```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2005  Compugen Ltd.


OM nucleic - protein search, using frame_plus_n2p model


Run on:    July 14, 2005, 17:06:19 ; Search time 168.5 Seconds
                                       (without alignments)
                                     4448.318 Million cell updates/sec

Title:          US-09-905-083A-30
Perfect score:  1780
Sequence:       1 ggattccgggctccatggc........aagaaacaacaaaccctcag 969

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters:        4211384

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09905083/runat_14072005_140352_10602/app_query.fasta_1.1159
-DB=A_Geneseq -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09905083 @CGN_1_1_224 @runat_14072005_140352_10602 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  A_Geneseq_16Dec04:*
         1: geneseqp1980s:*
         2: geneseqp1990s:*
         3: geneseqp2000s:*
         4: geneseqp2001s:*
         5: geneseqp2002s:*
         6: geneseqp2003as:*
         7: geneseqp2003bs:*
         8: geneseqp2004s:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID	Description
1	1364	76.6	253	2	AAR67888	Aar67888 Human str
2	1364	76.6	253	2	AAW05383	Aaw05383 Human amy
3	1364	76.6	253	5	ABB84421	Abb84421 Human SCC
4	1364	76.6	253	5	ABB84406	Abb84406 Human SCC
5	1364	76.6	253	5	AAU82740	Aau82740 Amino aci
6	1364	76.6	253	6	ABU07440	Abu07440 Protein d
7	1364	76.6	253	6	ABU07471	Abu07471 Protein d
8	1364	76.6	253	6	ABR58471	Abr58471 Human str
9	1364	76.6	253	7	ADB80484	Adb80484 Ovarian c
10	1364	76.6	253	7	ADJ68833	Adj68833 Human hea
11	1364	76.6	253	7	ADN39180	Adn39180 Cancer/an
12	1364	76.6	253	8	ADL06515	Adl06515 Human tum
13	1364	76.6	253	8	ADN04182	Adn04182 Antipsori
14	1364	76.6	253	6	ADR72880	Adr72880 Human ova
15	1350	75.8	250	6	ADA05732	Ada05732 Human NOV
16	1344	75.5	250	3	ADN62896	Adn62896 Human NOV
17	1337	75.1	257	3	AAB21326	Aab21326 Human HSC
18	1252	70.3	247	6	ADA05742	Ada05742 Human NOV
19	1252	70.3	247	8	ADN62906	Adn62906 Human NOV
20	1252	70.3	252	8	ADA05734	Ada05734 Human NOV
21	1252	70.3	252	8	ADN62898	Adn62898 Human NOV
22	1235	69.4	225	4	AAB98502	Aab98502 Human Str
23	1119	62.9	224	6	ADA05744	Ada05744 Human NOV
24	1119	62.9	224	6	ADN62908	Adn62908 Human NOV
25	1085.5	61.0	249	5	ABB84420	Abb84420 Porcine S
26	1035.5	58.2	198	8	ADA05736	Ada05736 Human NOV
27	1035.5	58.2	198	8	ADN62900	Adn62900 Human NOV
28	1011.5	56.8	249	5	ABB84423	Abb84423 Murine SC
29	996.5	56.0	243	5	ABB84419	Abb84419 Bovine SC
30	940.5	52.8	226	5	ABB84422	Abb84422 Rat SCCE
31	937.5	52.7	181	6	ADA05738	Ada05738 Human NOV
32	937.5	52.7	181	6	ADN62902	Adn62902 Human NOV
33	773	43.4	144	8	ADI39727	Adi39727 Stratum c
34	773	43.4	144	8	ADI37151	Adi37151 Stratum c
35	641.5	36.0	136	4	ABG23378	Abg23378 Novel hum
36	614	34.5	260	2	AAW10694	Aaw10694 Human rec
37	614	34.5	260	2	AAW12393	Aaw12393 Mouse neu
38	614	34.5	260	5	ABB57219	Abb57219 Mouse isc
39	614	34.5	260	5	ADI17073	Adi17073 Murine NO
40	614	34.5	260	8	ADI39731	Adi39731 Mouse neu
41	614	34.4	260	8	ADI37155	Adi37155 Mouse neu
42	612.5	34.4	293	2	AAY16777	Aay16777 Human ker
43	612.5	34.4	293	2	AAY30524	Aay30524 Human PDS
44	612.5	34.4	293	2	AAY38412	Aay38412 Human sec
45	612.5	34.4	293	2	AAY38426	Aay38426 Human sec

ALIGNMENTS

```
RESULT 1
AAR67888
ID   AAR67888 standard; protein; 253 AA.
XX
AC   AAR67888;
XX
DT   25-MAR-2003   (revised)
DT   09-AUG-1995   (first entry)
XX
DE   Human stratum corneum chymotropic recombinant enzyme (SCCE).
XX
KW   Stratum corneum chymotryptic enzyme; skin disorder; acne; psoriasis;
KW   callosities; keratosis pilaris; ichthyoses; eczema.
XX
OS   Homo sapiens.
XX
PN   WO9500651-A1.
XX
PD   05-JAN-1995.
XX
PF   20-JUN-1994;   94WO-IB000166.
XX
PR   18-JUN-1993;   93DK-00000725.
XX
PA   (SYMB-) SYMBICOM AB.
XX
PI   Egelrud T, Hansson L;
XX
DR   WPI; 1995-052088/07.
DR   N-PSDB; AAQ81203.
XX
PT   Nucleotide sequences encoding stratum corneum chymotryptic enzyme - and
PT   related vectors, transformed cells and polypeptides, useful for treating
PT   skin disorders, e.g. acne or psoriasis, and for identification of
```

```
PT   specific inhibitors.
XX
PS   Disclosure; Page 97; 137pp; English.
XX
CC   The enzyme encoded by this sequence is used in pharmaceutical, cosmetic
CC   and skin care products, especially to treat and prevent acne, xeroderma,
CC   or other hyperkeratotic conditions (e.g. callosities or keratosis
CC   pilaris), ichthyoses, psoriasis, eczema, etc. It is produced
CC   recombinantly following mammal, insect, plant, or microorganism
CC   transformation with plasmid pS507. (Updated on 25-MAR-2003 to correct PN
CC   field.)
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:              1.02e-134      Length:         253
Score:                  1364.00        Matches:        253
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            76.63%         Indels:         0
DB:                     2              Gaps:           0

US-09-905-083A-30 (1-969) x AAR67888 (1-253)

Qy    16 ATGGCAAGATCCCTTCTCCTGCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT   75
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr   20

Qy    76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGCGCCCCATGTGCAAGAGGCTCC  135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspAlaProCysAlaArgGlySer   40

Qy   136 CACCCATGGCAGTGGCCCTGCTCAGTGCAATCAGCTGCCACTGCGGGAGGCGTCCTCGGTC  195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal   60

Qy   196 AATGAGCCGCTGGGTGCTCACTGCGCGCCACTGCAAGATGAATGAGTACACCGTGCACCTG  255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AsnGluArgLeuTrpValThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu   80

Qy   256 GGCAGTGATACGCTGGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC  315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 GlySerAspThrLeuGlyArgGlnGluSerSerGluAspGlnGlyLeuGluValIleProArg   100

Qy   316 CACCCCGGCTACTCCACCACAGCCCATGTTAATGACCTCGTGTGAAGCTCAATAGC  375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 HisProGlyTyrSerThrThrHisValAsnAspLeuMetLeuValLysLeuAsnSer   120

Qy   376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTGCGCTGCGAACCCCCT  435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProCysSerGluProPro  140

Qy   436 GGAACCACCTGTACTGTCTCCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC  495
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 GlyThrThrCysThrValSerGlyThrTyrThrThrSerProAspValThrPhePro  160

Qy   496 TCTGACCTCATGTGCGTGGATGTCAAGCTCAAGCTGCGGGAGTTAC  555
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 SerAspLeuMetCysValAspValLysLeuSerProGlnAspCysThrLysValTyr  180

Qy   556 AAGGACTTACTGGAAATTCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAACGGCC  615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LysAspLeuLeuGluIleHisAlaValArgTrpHisProAspSerLysLysAsnAla  200

Qy   616 TGCAATGTGACTCAGGGGACACCGTTGGTGTGCAGAGTACCCTGCAAGGTCTGGTGTCC  675
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 CysAsnGlyAspSerGlyGlyProLeuValCysGlyArgGlyThrLeuGlnGlyLeuValSer  220

Qy   676 TGGGGAGACTTCCCTGCGGCCAACCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG  735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221 TrpGlyThrPheProCysGlyGlnProGlyValTyrThrGlnValCysLys  240

Qy   736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC  774
```

```
Db   241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg  253
         |||||||||||||||||||||||||||||||||||||||

RESULT 2
AAW05383
ID   AAW05383 standard; protein; 253 AA.
XX
AC   AAW05383;
XX
DT   31-DEC-1996  (first entry)
XX
DE   Human amyloid precursor protein protease.
XX
KW   Amyloid precursor protein protease; Alzheimer's disease; diagnosis;
KW   therapy.
XX
OS   Homo sapiens.
XX
PN   WO9631122-A1.
XX
PD   10-OCT-1996.
XX
PF   02-APR-1996;   96WO-US004294.
XX
PR   04-APR-1995;   95US-00416257.
XX
PA   (ELIL ) LILLY & CO ELI.
XX
PI   Dixon BP, Johnstone EM, Little SP;
XX
DR   WPI; 1996-464694/46.
DR   N-PSDB; AAT39783.
XX
PT   New isolated human amyloid precursor protein protease - used to develop
PT   prods. for the treatment or diagnosis of associated conditions, esp.
PT   Alzheimer's disease.
XX
PS   Claim 1; Page 44-45; 55pp; English.
XX
CC   Human amyloid precursor protein protease (AAW05383) is involved in the
CC   processing or clearance of amyloid precursor protein to form beta-amyloid
CC   peptide. Its amino acid sequence was deduced from a cDNA clone (AAT39783)
CC   obtd. from a human lung library. Recombinant protease can be produced in
CC   transformed or transfected prokaryotic (partic. E. coli) or eukaryotic
CC   (partic. AV-120 host cells. It is used to develop products for the design
CC   and testing of cpds. useful for treating or preventing conditions
CC   associated with beta-amyloid peptide, esp. Alzheimer's disease
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:              1.02e-134      Length:         253
Score:                  1364.00        Matches:        253
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            76.63%         Indels:         0
DB:                     2              Gaps:           0

US-09-905-083A-30 (1-969) x AAW05383 (1-253)

Qy    16 ATGGCAAGATCCCTTCTCCTGCCCTGCAGATCTTACTGCTATCCTTAGCCTTCGGAAACT   75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr   20

Qy    76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCC  135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspAlaProCysAlaArgGlySer   40

Qy   136 CACCCATGGCAGTGGCCCTGCTCAGTGCAATCAGCTGGCCACTGCGGGAGGCGGTCCTCGTC  195
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal   60

Qy   196 AATGAGCCGCTGGGTGCTCACTGCGCGCCACTGCAAGATGAATGAGTACACCGTGCACCTG  255
```

Alignment (continuation):

```
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AsnGluArgTrpValLeuThrArgAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu   80

Qy  256  GGCAGTGATACGCTGGGCACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC  315
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81  GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg  100

Qy  316  CACCCCGGCTACTCCACAGACCATGTTAATGACCCATGTCGTGAAGCTCAATAGC  375
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101  HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer  120

Qy  376  CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCT  435
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro  140

Qy  436  GGAACCACCTGACTGTCTCCGGCTGGGCACTACCACGAGCCCAGATGTGACCTTTCCC  495
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141  GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro  160

Qy  496  TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTAC  555
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161  SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr  180

Qy  556  AAGGACTTACTGGAAATTCATGCTGTGCGCTGGCATCCCGACTCCCAAGAAAAAACGCC  615
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LysAspLeuLeuGluIleHisAlaValArgTrpHisProAspSerLysLysLysArgAla  200

Qy  616  TGCAATGTGACTGGGGGGACCGTTGGTGTGCAGAGTACCCTGCAAGGTCTGGTGTCC  675
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201  CysAsnValThrGlyGlyThrValGlyValGlnSerThrLeuGlnGlyLeuValSer  220

Qy  676  TGGGGAACTTCCCTTGGGCCCAACCCAATGACCAGAGCTTACACTCAAGTGTGCAAG  735
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221  TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys  240

Qy  736  TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC  774
         |||||||||||||||||||||||||||||||||||||||
Db  241  PheThrLysTrpIleAsnAspThrMetLysLysHisArg  253
```

RESULT 3

ABB84421

ID   ABB84421 standard; peptide; 253 AA.
XX
AC   ABB84421;
XX
DT   08-NOV-2002 (first entry)
XX
DE   Human SCCE protein N-terminal fragment SEQ ID 48.
XX
KW   SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;
KW   serine protease; transgenic mammal; skin; skin disease; skin cancer;
KW   hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;
KW   pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO20026135-A2.
XX
PD   15-AUG-2002.
XX
PF   08-FEB-2002; 2002WO-IB001300.
XX
PR   09-FEB-2001; 2001CA-02332655.
PR   09-FEB-2001; 2001DK-0000218.
XX
PA   (EGEL/) EGELRUD T.
PA   (HANS/) HANSSON L.
XX
PI   Egelrud T, Hansson L;
XX
DR   WPI; 2002-643380/69.
XX
PT   Transgenic mammal or its embryo useful as model for human disease, has
PT   heterologous nucleotide sequence coding for stratum corneum chymotryptic
PT   enzyme operably linked to promoter that drives its expression in skin.
XX
PS   Example 6; Page 37; 74pp; English.
XX
CC   This invention describes a novel non-human transgenic mammal or mammalian
CC   embryo having integrated within its genome, a heterologous nucleotide
CC   sequence comprising at least a significant part of a nucleotide sequence
CC   coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,
CC   operably linked to a promoter that drives expression of heterologous scce
CC   or its variant in skin. The product of the invention is useful as a model
CC   for the study of disease with the aim of improving treatment, to relieve
CC   or ameliorate a pathogenic condition, for development or testing of a
CC   cosmetic or a pharmaceutical formulation, and for the development of a
CC   diagnostic method. It can also be used as a model for a skin disease or
CC   skin cancer. The invention is also useful for screening or identifying a
CC   compound or composition effective for the prevention or treatment of an
CC   abnormal or unwanted phenotype, and for screening or identifying a
CC   compound or composition effective for the prevention or treatment of
CC   inflammatory skin diseases selected from diseases consisting of epidermal
CC   hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,
CC   pruritus, atopic dermatitis, eczema, acne and inherited skin diseases
CC   with epidermal hyperkeratosis. The mammal of the invention is also useful
CC   as a model for further studies of itch mechanisms and the testing of
CC   potential compounds and compositions for relieve of various skin diseases
CC   where itch is a component. This sequence represents the N-terminal
CC   fragment of the human stratum corneum chymotryptic enzyme, SCCE
CC   synonymous with human kallikrein 7 (KLK7), used in the development of the
CC   transgenic mammals described in the invention
XX
SQ   Sequence 253 AA;

```
Alignment Scores:
Pred. No.:                                  Length:         253
Score:                 1.02e-134            Matches:        253
                       1364.00              Conservative:   0
Percent Similarity:    100.00%              Mismatches:     0
Best Local Similarity: 100.00%              Indels:         0
Query Match:           76.63%               Gaps:           0
DB:                    5
```

```
US-09-905-083A-30 (1-969) x ABB84421 (1-253)

Qy   16  ATGGCAAGATCCCTTCCTGCCCCTGCCCGAGATCTTACTGCTATCCTTAGCCTTGGAAACT  75
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MetAlaArgSerLeuLeuProGlnIleLeuLeuLeuLeuSerLeuAlaLeuGluThr  20

Qy   76  GCAGGAGGAGAGCCCAGGGTGACAAGATTATTGATGAGGCCCATGTGCAAGAGGCTCC  135
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21  AlaGlyGluGluGlnAlaGlnGlyAspLysLysIleIleAspGluAlaProCysAlaArgGlySer  40

Qy  136  CACCCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGTGCCATCAGCTCCCACTGCGGAGGCGTCCTGGTC  195
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41  HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal  60

Qy  196  AATGAGCGCTGGGTGCTCACTCGCGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG  255
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AsnGluArgTrpValLeuThrArgAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu  80

Qy  256  GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC  315
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81  GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg  100

Qy  316  CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCATGTTGGTGAAGCTCAATAGC  375
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101  HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer  120

Qy  376  CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCT  435
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro  140

Qy  436  GGAACCACCTGACTGTCTCCGGCTGGGC  495
         ||||||||||||||||||||||||||||
Db  141  GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro  160
```

RESULT 4
ABB84406
ID   ABB84406 standard; protein; 253 AA.
XX
AC   ABB84406;
XX
DT   08-NOV-2002 (first entry)
XX
DE   Human SCCE protein.
XX
KW   SCCE; human; stratum corneum chymotryptic enzyme; kallikrein 7;
KW   serine protease; transgenic mammal; skin; skin cancer;
KW   hyperkeratosis; acanthosis; epidermal inflammation; dermal inflammation;
KW   pruritus; atopic dermatitis; eczema; acne; itch; KLK7; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO200262135-A2.
XX
PD   15-AUG-2002.
XX
PF   08-FEB-2002; 2002WO-IB001300.
XX
PR   09-FEB-2001; 2001CA-02332655.
PR   09-FEB-2001; 2001DK-00000218.
XX
PA   (EGEL/) EGELRUD T.
PA   (HANS/) HANSSON L.
XX
PI   Egelrud T, Hansson L;
XX
DR   WPI; 2002-643380/69.
DR   N-PSDB; ABQ76226.
XX
PT   Transgenic mammal or its embryo useful as model for human disease, has
PT   heterologous nucleotide sequence coding for stratum corneum chymotryptic
PT   enzyme operably linked to promoter that drives its expression in skin.
XX
PS   Claim 10; Page 58-59; 74pp; English.
XX
CC   This invention describes a novel non-human transgenic mammal or mammalian
CC   embryo having integrated within its genome, a heterologous nucleotide
CC   sequence comprising at least a significant part of a nucleotide sequence
CC   coding for a stratum corneum chymotryptic enzyme (SCCE) or its variant,
CC   operably linked to a promoter that drives expression of heterologous scce
CC   or its variant in skin. The product of the invention is useful as a model
CC   for the study of disease with the aim of improving treatment, to relieve
CC   or ameliorate a pathogenic condition, for development or testing of a
CC   cosmetic or a pharmaceutical formulation, and for the development of a
CC   diagnostic method. It can also be used as a model for a skin disease or
CC   skin cancer. The invention is also useful for screening or identifying a
CC   compound or composition effective for the prevention or treatment of an
CC   abnormal or unwanted phenotype, and for screening or identifying a
CC   compound or composition effective for the prevention or treatment of
CC   inflammatory skin diseases selected from diseases consisting of epidermal
CC   hyperkeratosis, acanthosis, epidermal inflammation, dermal inflammation,
CC   pruritus, atopic dermatitis, eczema, acne and inherited skin diseases
CC   with epidermal hyperkeratosis. The mammal of the invention is also useful
CC   as a model for further studies of itch mechanisms and the testing of
CC   potential compounds and compositions for relieve of various skin diseases
CC   where itch is a component. This sequence represents the human stratum
CC   corneum chymotryptic enzyme, SCCE which is a serine protease synonymous
CC   with human kallikrein 7 (KLK7) and is used in the development of the
CC   transgenic mammals described in the invention
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:   Length: 253
Score: 1364.00   Matches: 253
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 76.63%   Indels: 0
DB: Gaps: 0

US-09-905-083A-30 (1-969) x ABB84406 (1-253)

```
Qy   16 ATGGCAAGATCCCTTCCTGCCCCTGCGAGATCTTACTGCTATCCTTAGCCTTGGAAACT   75
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MetAlaArgSerLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr      20

Qy   76 GCAGGAGAAGAGCCCAGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCC    135
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer   40

Qy  136 CACCCCATGGCAGTGGCCCTGCCAGTGGCAATCAGCTTCAGTGCTGGGAGGCGTCCTGGTC  195
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal   60

Qy  196 AATGAGCGCTGGGTGCTCACTGCGCCACTGCAAGATGAATGAGTACACCGTGCACCTG    255
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu   80

Qy  256 GGCAGTGATACGCTGGGCACAGGAGAGCTCAGAGGATCAAGGCCTGAAGTCATTCCGC    315
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg  100

Qy  316 CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC  375
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer  120

Qy  376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGGCTGCGAACCCCT  435
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro  140

Qy  436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCAGCCCAGATGACCTTTCCC      495
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspAspLeuSerPhePro  160

Qy  496 TCTGACCTCATGTGCGTGGATGTGCAAGCTCATCTCCCCCAGGACTGACGAGGTTTAC    555
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr  180

Qy  556 AAGGACTTACTGGAAAATTCCATGCTCGTGCGCTGGCATCCCGACTCCAAGAAAAACGCC  615
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LysAspLeuLeuGluAsnSerMetLeuValArgTrpHisProAspSerLysLysAsnAla  200

Qy  616 TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGTACCCTGCAAGGTCTGTGTCC    675
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyValSer     220

Qy  676 TGGGGAACTTCCCTTGCGGCCAACCCAATGACCCAGAGTCTACACTCAAGTGTGCAAG    735
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys  240

Qy  736 TTCCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC                      774
         ||||||||||||||||||||||||||||||||||||||||
Db  241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg                       253
```

Qy   736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
           |||||||||||||||||||||||||||||||||||||||
Db   241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 5
AAU82740
ID   AAU82740 standard; protein; 253 AA.
XX
AC   AAU82740;
XX
DT   23-APR-2002 (first entry)
XX
DE   Amino acid sequence of novel human protease #39.
XX
KW   Human; protease; cancer; immune-related disorder; cardiovascular disease;
KW   neuronal-associated disease; metabolic disorder; inflammatory disorder;
KW   nervous system disorder; sexual dysfunction; pain; mood disorder;
KW   hypertension; psychotic disorder; neurological disorder; dyskinesia;
KW   viral infection; human immunodeficiency virus; HIV; non-viral infection;
KW   ocular disease; cytostatic; enzyme.
XX
OS   Homo sapiens.
XX
PN   WO20200860-A2.
XX
PD   03-JAN-2002.
XX
PF   26-JUN-2001; 2001WO-US020171.
XX
PR   26-JUN-2000; 2000US-0214047P.
XX
PA   (SUGB-) SUGEN INC.
XX
PI   Plowman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;
PI   Charydczak G;
XX
DR   WPI; 2002-139913/18.
DR   N-PSDB; ABK31782.
XX
PT   Nucleic acids encoding novel human proteases, useful for useful for
PT   treating diseases and disorders such as cancers, immune-related diseases
PT   and disorders, cardiovascular disease (e.g. restenosis) and inflammatory
PT   disorders.
XX
PS   Claim 6; Fig 2N; 313pp; English.
XX
CC   The present invention relates to the isolation of novel human proteases,
CC   and the nucleic acids encoding them. The sequences of the invention are
CC   useful for treating diseases and disorders such as cancers (e.g. breast,
CC   colon, lung), immune-related diseases and disorders (e.g. inflammatory
CC   diseases and asthma), cardiovascular diseases (e.g. restenosis and
CC   coronary thrombosis), brain or neuronal-associated diseases, metabolic
CC   disorders (e.g. diabetes, obesity), inflammatory disorders (e.g.
CC   rheumatoid arthritis and psoriasis), central or peripheral nervous system
CC   diseases, migraines, pain, sexual dysfunction, mood disorders, attention
CC   disorders, cognition disorders, hypotension, hypertension, psychotic
CC   disorders, neurological disorders (e.g. Alzheimer's disease, Parkinson's
CC   disease) and dyskinesias. The nucleic acids and polypeptides are also
CC   useful for treating viral infections caused by human immunodeficiency
CC   virus (HIV), and non-viral infections such as ocular disease (e.g.
CC   glaucoma) and macular degeneration. AAU82702-AAU82760 represent the novel
CC   human proteases of the invention
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:   Length: 253
Score: 1364.00   Matches: 253
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 76.63%   Indels: 0
DB: 5   Gaps: 0

US-09-905-083A-30 (1-969) x AAU82740 (1-253)

Qy   16 ATGGCAAGATCCCTTCCTGCCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT   75
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MetAlaArgSerLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr   20

Qy   76 GCAGGAGAAGCCCAGGTGACAAGATTATTGATGGGCGCCCATGTGCAAGAGGCTCC 135
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AlaGlyGluGluAlaGlnValThrArgLeuLeuIleGlyAlaProCysAlaArgGlySer   40

Qy   136 CACCCATGCAGGTGGCCCTGCTCAGTGGCAATCAGTCTGCTGCGGAGGCGTCTGGTC 195
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuValHisCysGlyValLeuVal   60

Qy   196 AATGAGGCTGGTGCTCACTGCCGCCACTGCAAGATGAATGAGTACCGTGCACCTG 255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu   80

Qy   256 GGCAGTGATACGGTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTCAAGTCATTCCGC 315
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg 100

Qy   316 CACCCGGCTACTCCACACAGACCCATGTAATGACCTCGTCGTGAAGCTCAATAGC 375
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuValLeuLysLeuAsnSer 120

Qy   376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCGCCTCCGCTGCGAACCCCCT 435
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuArgLeuProSerArgCysGluProPro 140

Qy   436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACAGAGCCCAGATGTGACCTTTCCC 495
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160

Qy   496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCAGGACTGCACGAAGGTTTAC 555
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180

Qy   556 AAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCCAAGAAAAACGCC 615
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProThrProLysLysLysAsnAla 200

Qy   616 TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCC 675
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy   676 TGGGGAACTTTCCCTGCGGCCAACCCAATGACACCAGAGTCTACACTGACCAAGTGTGCAAG 735
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy   736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
           |||||||||||||||||||||||||||||||||||||||
Db   241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 6
ABU07440
ID   ABU07440 standard; protein; 253 AA.
XX
AC   ABU07440;
XX
DT   28-JAN-2003 (first entry)
XX
DE   Protein differentially regulated in prostate cancer #43.
XX
KW   Prostate cancer; gene expression; differential regulation;
KW   molecular marker; drug target; cancer detection; cancer diagnosis;
KW   cancer staging; cancer grading; cancer assessing; cancer monitoring.
XX
OS   Homo sapiens.
XX
PN   WO20281638-A2.
XX
PD   17-OCT-2002.

```
XX
PF   08-APR-2002; 2002WO-US010824.
XX
PR   06-APR-2001; 2001US-0281731P.
PR   06-APR-2001; 2001US-0281732P.
XX
PA   (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI   Sun Z, Jay G;
XX
DR   WPI; 2003-058520/05.
DR   N-PSDB; ABX10343.
XX
PT   Novel genes which are differentially regulated in prostate cancer, useful
PT   for diagnosing prostate cancer in prostate tissue sample and assessing
PT   therapeutic or preventive intervention in prostate cancer patients.
XX
PS   Claim 1; Page 293-294; 416pp; English.
XX
CC   The invention describes genes (I) which are differentially regulated in
CC   prostate cancer. (I) is useful for diagnosing a prostate cancer in a
CC   sample comprising prostate tissue, which involves determining the number
CC   of target genes which are differentially-regulated in the sample, where
CC   the number is indicative of the probability that the sample comprises
CC   prostate cancer. (I) is useful for assessing a therapeutic or preventive
CC   intervention in a subject having a prostate cancer, which involves
CC   determining the expression levels in a sample comprising prostate tissue
CC   of target genes which are differentially-regulated in prostate cancer.
CC   Preferably, the expression levels of at least 10 genes are determined.
CC   (I) is also useful for identifying agents that modulate a biological
CC   activity of a polypeptide differentially-regulated in prostate cancer
CC   cells, which involves contacting a polypeptide differentially-regulated
CC   in prostate cancer cells with a test agent under conditions effective for
CC   the test agent to modulate a biological activity of the polypeptide, and
CC   determining whether the test agent modulates the biological activity. (I)
CC   is useful as molecular markers, as drug targets, and for detecting,
CC   diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC   preventing or treating, determining predisposition to diseases and
CC   conditions especially relating to prostate cancer. (I) and its expression
CC   products are used in the diagnostic test to assay for presence of cancer
CC   e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC   blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC   of cancer, its stage of development, the nature of genetic defect, etc.
CC   The polypeptide encoded by (I) can be used as target for therapy or drug
CC   discovery. (I) can also be used for expressing the polypeptide and thus
CC   for searching specific binding partners of the polypeptide. The identification
CC   of specific genes, and groups of genes, expressed in pathways
CC   physiologically relevant to prostate cancer permits the definition of
CC   functional and disease pathways and the delineation of targets in these
CC   pathways which are useful in diagnostic, therapeutic, and clinical
CC   applications. This is the amino acid sequence of a protein differentially
CC   regulated in prostate cancer
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:                  1.02e-134     Length:          253
Score:                      1364.00       Matches:         253
Percent Similarity:         100.00%       Conservative:    0
Best Local Similarity:      100.00%       Mismatches:      0
Query Match:                76.63%        Indels:          0
DB:                         6             Gaps:            0

US-09-905-083A-30 (1-969) x ABU07440 (1-253)

Qy     16 ATGGCAAGATCCCTTCTCCTGCGCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy     76 GCAGGAGAAGAAGCCCAGGGTGACAAGATTATTGATGCGCGCCCCATGTGCAAGAGGCTCC 135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspAlaArgProCysAlaArgGlySer 40
```

```
Qy    136 CACCCATGGCAGGTGGGCCCTGCTCCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 HisProGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal 60

Qy    196 AATGAGCGGCTGGGTGCTCACTGCGCCCCACTGCAAGATGAATGAGTACACCGTGCACCTG 255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AsnGluArgTrpValLeuThrAlaAlaHisCysAlaAsnGluTyrThrValHisLeu 80

Qy    256 GGCAGTGATACGCTGGGCGACACGAGAGCTCAGGAGATCAAGGGCCTCAAGTCATTCCGC 315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg 100

Qy    316 CACCCGGCTACTCCACCAGACCCATGTTAATGACCTCGTGAAGCTCAATAGC 375
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy    376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGACCCCCT 435
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy    436 GGAACCACCTGTACTGCTCCGGCTGGGGCACTACCACCAGCCCAGATGTGACCTTTCCC 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160

Qy    496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCCAGGACTGCACGAAGGTTTAC 555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 SerAspLeuMetCysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyr 180

Qy    556 AAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCCAAGAAAAAACGCC 615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProThrProLysLysLysAsnAla 200

Qy    616 TGCAATGTGACTCAGGGGGACCGTTGCTGTGTGCAGAGTACCCTGCAAGGTCTGGTGTCC 675
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy    676 TGGGGAACTTCCCTGCGGCCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy    736 TTCACCAAGTGGATAAATGACACACATGAAAAAGCATCGC 774
          ||||||||||||||||||||||||||||||||||||||||
Db    241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

```
RESULT 7
ABU07471
ID . ABU07471 standard; protein; 253 AA.
XX
AC   ABU07471;
XX
DT   28-JAN-2003 (first entry)
XX
DE   Protein differentially regulated in prostate cancer #74.
XX
KW   Prostate cancer; gene expression; differential regulation;
KW   molecular marker; drug target; cancer detection; cancer diagnosis;
KW   cancer staging; cancer grading; cancer assessing; cancer monitoring.
XX
OS   Homo sapiens.
XX
PN   WO20281638-A2.
XX
PD   17-OCT-2002.
XX
PF   08-APR-2002; 2002WO-US010824.
XX
PR   06-APR-2001; 2001US-0281731P.
PR   06-APR-2001; 2001US-0281732P.
XX
PA   (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
PI   Sun Z, Jay G;
```

```
Qy  256 GGCAGTGATACGCTGGGCGACAGGAGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC 315
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg 100

Qy  316 CACCCCGGCTACTCCACACAGACCCCATGTAATGACCTGCTGAGCTCAATAGC 375
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy  376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCGCTGCGAACCCCT 435
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GlnAlaArgLeuSerSerMetValLysSerValArgLeuProSerArgCysGluProPro 140

Qy  436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC 495
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160

Qy  496 TCTGACCTCATGGTGCGTGGATGTCAAGCTCATCTCCCCCCAGGACTGCACGAAGTTTAC 555
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161 SerAspLeuMetCysValAlaAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180

Qy  556 AAGGACTTACTGGAAAATTCCATGCTGTCGCGTGCAGACTCCGACTCCAAGAAAAACGCG 615
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200

Qy  616 TGCAATGGTGACTCAGGGGACCGTTGGTGTCGAGAGGTACCCTGCAAGGTCGTGTCC 675
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy  676 TGGGGAGAACTTCCCTGCGGCCACCCAATGACCGCCAGGAGTCTACACTCCAAGTGTGCAAG 735
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy  736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
            ||||||||||||||||||||||||||||||||||||||
Db  241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

```
RESULT 8
ABR58471
ID   ABR58471 standard; protein; 253 AA.
XX
AC   ABR58471;
XX
DT   07-JUL-2003   (first entry)
XX
DE   Human stratum corneum chymotryptic enzyme - ovarian cancer clone O1676P.
XX
KW   Human; cytostatic; gene therapy; vaccine; cancer; ovarian cancer.
XX
OS   Homo sapiens.
XX
PN   WO200302946B-A1.
XX
PD   10-APR-2003.
XX
PF   02-OCT-2002; 2002WO-US031467.
XX
PR   02-OCT-2001; 2001US-0327135P.
PR   30-MAY-2002; 2002US-0384531P.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Algate PA, Mannion J;
XX
DR   WPI; 2003-372001/35.
XX
PT   New polynucleotide and polypeptide useful for diagnosing and/or treating
PT   cancer, particularly ovarian cancer, and as a vaccine.
XX
PS   Claim 2; Page 157-158; 169pp; English.
XX
CC   The invention relates to a novel isolated polynucleotide. The
CC   polynucleotides of the invention have cytostatic activity, and may have a
CC   use in gene therapy, and in a vaccine. The composition and methods are
```

```
XX
DR   WPI; 2003-058520/05.
DR   N-PSDB; ABX10375.
XX
PT   Novel genes which are differentially regulated in prostate cancer, useful
PT   for diagnosing prostate cancer in prostate tissue sample and assessing
PT   therapeutic or preventive intervention in prostate cancer patients.
XX
PS   Claim 1; Page 351; 416pp; English.
XX
CC   The invention describes genes (I) which are differentially regulated in
CC   prostate cancer. (I) Is useful for diagnosing a prostate cancer in a
CC   sample comprising prostate tissue, which involves determining the number
CC   of target genes which are differentially-regulated in the sample, where
CC   the number is indicative of the probability that the sample comprises
CC   prostate cancer. (I) Is useful for assessing a therapeutic or preventive
CC   intervention in a subject having a prostate cancer, which involves
CC   determining the expression levels in a sample comprising prostate tissue
CC   of target genes which are differentially-regulated in prostate cancer.
CC   Preferably, the expression levels of at least 10 genes are determined.
CC   (I) is also useful for identifying agents that modulate a biological
CC   activity of a polypeptide differentially-regulated in prostate cancer
CC   cells, which involves contacting a polypeptide differentially-regulated
CC   in prostate cancer cells with a test agent under conditions effective for
CC   the test agent to modulate a biological activity of the polypeptide, and
CC   determining whether the test agent modulates the biological activity. (I)
CC   is useful as molecular markers, as drug targets, and for detecting,
CC   diagnosing, staging, grading, assessing, monitoring, prognosticating,
CC   preventing or treating, determining predisposition to diseases and
CC   conditions especially relating to prostate cancer. (I) and its expression
CC   products are used in the diagnostic test to assay for presence of cancer
CC   e.g., in tissue sections, in biopsy sample, in total RNA, in lymph, in
CC   blood etc. (I) is useful for assessing cancer e.g., to determine the type
CC   of cancer, its stage of development, the nature of genetic defect, etc.
CC   The polypeptide encoded by (I) can be used as target for therapy or drug
CC   discovery, (I) can also be used for expressing the polypeptide and thus
CC   for searching specific binding partners of the polypeptide. (I) is useful
CC   in therapeutic applications to treat prostate cancer. The identification
CC   of specific genes, and groups of genes, expressed in pathways
CC   physiologically relevant to prostate cancer permits the definition of
CC   functional and disease pathways and the delineation of targets in these
CC   pathways which are useful in diagnostic, therapeutic, and clinical
CC   applications. This is the amino acid sequence of a protein diferentially
CC   regulated in prostate cancer
XX
SQ   Sequence 253 AA;
```

```
Alignment Scores:
Pred. No.:              1.02e-134        Length:          253
Score:                  1364.00          Matches:         253
Percent Similarity:     100.00%          Conservative:    0
Best Local Similarity:  100.00%          Mismatches:      0
Query Match:            76.63%           Indels:          0
DB:                     6                Gaps:            0
```

```
US-09-905-083A-30 (1-969) x ABU07471 (1-253)

Qy   16 ATGGCAAGATCCCTTCTCCGCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MetAlaArgSerLeuLeuLeuProLeuLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy   76 GCAGGAGAAGAGCCCAGGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGGCTCC 135
            |||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AlaGlyGluGluAlaGlnGlyAspLysLysIleIleAspGlyAlaProCysAlaArgGlySer 40

Qy  136 CACCCATGCAGGCTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGTC 195
            ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisLeuHisCysGlyGlyValLeuVal 60

Qy  196 AATGAGCGGCTGGGTGCTCACTGCCCCCACTGCAAGATGAATGAGTACACCGTGCACCTG 255
            |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AsnGluArgLeuThrLeuThrAlaAlaHisCysLysMetAsnGlnLeuTyrThrValHisLeu 80
```

CC   useful in diagnosing and/or treating cancer, particularly ovarian cancer.
CC   The composition may also be used as a vaccine to prevent cancer. The
CC   present sequence is used in the exemplification of the invention
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:   1.02e-134   Length: 253
Score: 1364.00   Matches: 253
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 76.63%   Indels: 0
DB: 6   Gaps: 0

US-09-905-083A-30 (1-969) x ABR58471 (1-253)

```
Qy    16  ATGGCAAGATCCCTTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT  75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MetAlaArgSerLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr  20

Qy    76  GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCC  135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21  AlaGlyGluGluAlaGlnValAspLysIleIleIleAspGlyAlaProCysAlaArgGlySer  40

Qy   136  CACCCATGCAGGTGGCCCTGCTCAGTGGCAATCAGCTGCCACTGCGAGGCGTCCTGGTC  195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41  HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal  60

Qy   196  AATGAGCCTGGGTGCTCACTGCCCCACTGCAAGATGAATGAGTACACCGTGCACCTG  255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61  AsnGluArgTrpValLeuThrLeuAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu  80

Qy   256  GGCAGTGATGACGCTGGGCACAGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC  315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81  GlySerAspThrLeuGlyHisArgArgAlaGlnArgIleLysAlaSerLysSerPheArg  100

Qy   316  CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCGTGAAGTCAATAGC  375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101  HisProGlyTyrSerThrGlnThrHisValAsnAspLeuValLysLeuAsnSer  120

Qy   376  CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGTGCCCTGCGAACCCCCT  435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GlnAlaArgLeuSerSerMetValLysLysValArgCysProCysGluProPro  140

Qy   436  GGAACCACCTGTACTGTCTCCGGCTGGGGCACTCCACCAGCCCAGATGTGACCTTTCCC  495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141  GlyThrThrCysThrValSerGlyTrpGlyThrProThrSerProAspValThrPhePro  160

Qy   496  TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCCACGAAGGTTTAC  555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161  SerAspLeuMetCysValAspValLysLeuIleSerProArgThrAlaThrLysValTyr  180

Qy   556  AAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAACGCC  615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProThrProSerLysLysAsnAla  200

Qy   616  TGCAATGTGACTCAGGGGGCCGTTGTGTGTCGAGAGTACCCGTCGAAGGTCTCGGTCC  675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201  CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer  220

Qy   676  TGGGGAACTTTCCCTTGCGGCACCCCAAGCCCAAATGACACCCAGGAGTCTACACTCAAGTGTGCAAG  735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221  TrpGlyThrPhePheProCysGlyThrGlnProAsnAspThrGlnGluSerThrGlnValCysLys  240

Qy   736  TTCACCAAGTGGATAAATGACCACCATGAAAAAAGCATCGC  774
          |||||||||||||||||||||||||||||||||||||||||
Db   241  PheThrLysTrpIleAsnAspThrMetLysLysHisArg  253
```

RESULT 9
ADB80484
ID   ADB80484   standard; protein; 253 AA.
XX
AC   ADB80484;

XX
DT   04-DEC-2003   (first entry)
XX
DE   Ovarian cancer-associated protein #24.
XX
KW   cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW   post-operative chemotherapy; radiation therapy; tumour prognosis;
KW   pre-cancerous lesion detection.
XX
OS   Homo sapiens.
XX
PN   WO200210223S-A2.
XX
PD   27-DEC-2002.
XX
PF   18-JUN-2002; 2002WO-US019297.
XX
PR   18-JUN-2001; 2001US-0299234P.
PR   27-AUG-2001; 2001US-0315287P.
PR   05-SEP-2001; 2001US-0317544P.
PR   13-NOV-2001; 2001US-0350666P.
PR   12-APR-2002; 2002US-0372246P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Mack DH,   Gish KC;
XX
DR   WPI; 2003-167431/16.
DR   N-PSDB; ADB80483.
XX
PT   Detecting an ovarian cancer-associated transcript in a cell from a
PT   patient, comprises contacting a biological sample from the patient with a
PT   polynucleotide that hybridizes to an ovarian cancer gene.
XX
PS   Claim 13; Page 291; 332pp; English.
XX
CC   The invention relates to a method of detecting an ovarian cancer-
CC   associated transcript in a cell from a patient, by contacting a
CC   biological sample from the patient with a polynucleotide that selectively
CC   hybridizes to a sequence at least 80% identical to any of one of 80
CC   nucleic acid sequences given in the specification. The method is useful
CC   in diagnosing ovarian cancer and in identifying and using agents and/or
CC   targets that inhibit ovarian cancer. The nucleic acid molecule,
CC   polypeptide and the antibody may also be used in detecting ovarian
CC   cancers, monitoring and early detection of relapse following treatment,
CC   monitoring response to therapy, selecting patients for post-operative
CC   chemotherapy or radiation therapy, in selecting mode of therapy,
CC   determining tumour prognosis, early detection of pre-cancerous lesions,
CC   and as vaccines. This sequence corresponds to one of the proteins used
CC   for the detection method of the invention.
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:   1.02e-134   Length: 253
Score: 1364.00   Matches: 253
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 76.63%   Indels: 0
DB: 7   Gaps: 0

US-09-905-083A-30 (1-969) x ADB80484 (1-253)

```
Qy    16  ATGGCAAGATCCCTTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT  75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1  MetAlaArgSerLeuLeuProLeuGlnIleIleLeuLeuLeuSerLeuAlaLeuGluThr  20

Qy    76  GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCC  135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    21  AlaGlyGluGluAlaGlnValAspLysIleIleIleGluAspGlyAlaProCysAlaArgGlySer  40

Qy   136  CACCCATGCAGGTGGCCCTGCTCAGTGGCAATCAGCTGCCACTGCGAGGCGTCCTGGTC  195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
RESULT 10
ADJ68833
ID   ADJ68833 standard; protein; 253 AA.
XX
AC   ADJ68833;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Human heart mitochondrial protein as a therapeutic target SeqID639.
XX
KW   mitochondrial; human; screening assay; diabetes mellitus;
KW   Huntington's disease; osteoarthritis;
KW   Leber's hereditary optic neuropathy; LHON;
KW   mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW   myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW   neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW   osteopathic; ophthalmological; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2003087768-A2.
XX
PD   23-OCT-2003.
XX
PF   04-APR-2003; 2003WO-US010870.
XX
PR   12-APR-2002; 2002US-0372843P.
PR   17-JUN-2002; 2002US-0389987P.
PR   20-SEP-2002; 2002US-0412418P.
XX
PA   (MITO-) MITOKOR.
PA   (BUCK-) BUCK INST AGE RES.
XX
PI   Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI   Warnock DE;
XX
DR   WPI; 2003-845369/78.
XX
PT   Identifying a mitochondrial target for drug screening assays and for
PT   treating diseases associated with altered mitochondrial function,
PT   comprises detecting a modified polypeptide in a sample and correlating
PT   with the disease.
XX
PS   Claim 1; SEQ ID NO 639; 180pp; English.
XX
CC   This invention relates to novel mitochondrial targets that can be used
CC   for therapeutic intervention in treating a disease associated with
CC   altered mitochondrial function. Specifically, it refers to a method for
CC   identifying proteins of the human heart mitochondrial proteome that are
CC   useful for drug screening assays, as well as therapeutic targets. The
CC   present invention describes a method for identifying such proteins that
CC   can be used in the treatment of various diseases associated with altered
CC   mitochondrial function including diabetes mellitus, Huntington's disease,
CC   osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC   encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC   ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC   compositions have neuroprotective, nootropic, antidiabetic,
CC   anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC   cytostatic activities. This polypeptide sequence is a human heart
CC   mitochondrial protein of the invention.
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:                          Length:            253
Score:                 1.02e-134    Matches:           253
Percent Similarity:    1364.00      Conservative:      0
Best Local Similarity: 100.00%      Mismatches:        0
Query Match:           76.63%       Indels:            0
DB:                                 Gaps:              0

US-09-905-083A-30 (1-969) x ADJ68833 (1-253)

Qy   16 ATGGCAAGATCCCTTCCTCCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGAAACT 75
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy   76 GCAGGAGGAAGAAGCCCAGGTGACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCC 135
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40

Qy  136 CACCCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy  196 AATGAGCGGCTGGGTGCTCACTGCGCCACCCAGGAGTCTACACTGCACCTG 255
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AsnGluArgTrpValLeuThrAlaAlaHisCysGlyThrValHisLeu 80

Qy  256 GGCAGTAGACGCTGGGGCACAGAGAGCTCAGGAGATCAAGGCCTCGAAGTCATTCCGC 315
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 GlySerAspThrLeuGlyThrGluSerSerGlyAspGlnArgAlaSerLysSerPheArg 100

Qy  316 CACCCCGGCTACTCCACCAGACCCATGTTAATGACCTGTCGTGAAGCTCAATAGC 375
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101 HisProGlyTyrSerThrArgGlnThrHisValAsnAspLeuMetLeuValLysAsnSer 120

Qy  376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCCGCTGGGAACCCCCT 435
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerProLeuGlyThr 140

Qy  436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCAGAGCCCAGATGTGACCTTTCCC 495
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160

Qy  496 TCTGACCTCATGTGCGTGATGTCAAGCTCATCTCCCCAGGACTGCACGGAAGTTTAC 555
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161 SerAspLeuMetCysValMetSerSerSerProGlnAspCysThrLysValTyr 180

Qy  556 AACGGACTTACTGGAAAATTCCATGCTGTGCCTGGCATCCCGACTCCCAAGAAAAACGCC 615
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysAsnAla 200

Qy  616 TGCAATGGTGACTCAGGGGGACCCGTTGGTGTGCAGAGGTACCCTGCAAGGTYTGGTGTCC 675
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy  676 TGGGGAACTTCCCTTGCGGCCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221 TrpGlyThrPheProCysGlyGlnProAsnAspProMetThrGlnValCysLys 240

Qy  736 TTCACCAAGTGGATAAATGACACCATGAAAAAAGCATCGC 774
         ||||||||||||||||||||||||||||||||||||||||
Db  241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

```
Qy   496  TCTGACCTCATGTGCGTGGATGTCAAGTCTCATCTCCCCCAGAGACTGCACGAGAGGTTTAC  555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161  SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr    180

Qy   556  AAGGACTTACTGGAAAATTCCATGTGCGTGGCTGGCATCCCCGACTCCAAGAAAACGCC     615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181  LysAspLeuLeuGluAsnSerMetCysValAlaGlyIleProAspSerLysLysAsnAla   200

Qy   616  TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCTGCAAGGTCTGGTGTCC    675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201  CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer   220

Qy   676  TGGGGAACTTTCCCTTGGCGGCCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG   735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221  TrpGlyThrPhePheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys  240

Qy   736  TTCACCAAGTGGATAAATGACACCACCATGAAAAAGCATCGC   774
          ||||||||||||||||||||||||||||||||||||
Db   241  PheThrLysTrpIleAsnAspThrThrMetLysLysHisArg  253
```

RESULT 11

```
ADN39180
ID   ADN39180 standard; protein; 253 AA.
XX
AC   ADN39180;
XX
DT   17-JUN-2004  (first entry)
XX
DE   Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:498.
XX
OS   Homo sapiens.
XX
KW   Human; differential expression; cancer; angiogenic disorder;
KW   fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW   inflammatory disease; autoimmune disease;
KW   retinal neovascularistaion syndrome; scarring; uterine fibroid;
KW   detection; diagnosis; prognosis; drug screening; drug targeting;
KW   wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW   vulnerary; gene therapy; vaccine.
XX
PN   WO200342661-A2.
XX
PD   22-MAY-2003.
XX
PF   13-NOV-2002; 2002WO-US036810.
XX
PR   13-NOV-2001; 2001US-0350666P.
PR   21-NOV-2001; 2001US-0332464P.
PR   29-NOV-2001; 2001US-0334393P.
PR   03-DEC-2001; 2001US-0335394P.
PR   14-DEC-2001; 2001US-0340376P.
PR   08-JAN-2002; 2002US-0347211P.
PR   10-JAN-2002; 2002US-0347349P.
PR   08-FEB-2002; 2002US-0355250P.
PR   13-FEB-2002; 2002US-0356714P.
PR   20-FEB-2002; 2002US-0359077P.
PR   29-MAR-2002; 2002US-0368809P.
PR   04-APR-2002; 2002US-0370110P.
PR   12-APR-2002; 2002US-0372246P.
PR   05-JUN-2002; 2002US-0386614P.
PR   16-JUL-2002; 2002US-0396839P.
PR   22-JUL-2002; 2002US-0397775P.
PR   22-JUL-2002; 2002US-0397845P.
PR   09-SEP-2002; 2002US-0409450P.
XX
PA   (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI   Afar D, Aziz N, Ginsburg WM, Gish KC, Glynne R, Hevezi PA;
PI   Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR   WPI; 2003-468649/44.
DR   N-PSDB; ADN39179.
```

```
XX
PT   Determining the presence or absence of a pathological cell in a patient,
PT   useful for diagnosing, prognosing or treating cancer, comprises detecting
PT   a nucleic acid in a biological sample.
XX
PS   Claim 12; SEQ ID NO 498; 1385pp; English.
XX
CC   The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC   whose expression is upregulated or downregulated in specific cancers or
CC   other diseases such as angiogenic or fibrotic disorders, and to methods
CC   of determining the presence or absence of a pathological cell in a
CC   patient by detecting a nucleic acid at least 80% identical to those of
CC   the invention or by detecting a polypeptide of the invention. The
CC   invention also relates to expression vectors and host cells comprising a
CC   nucleic acid of the invention; antibodies which specifically bind a
CC   polypeptide of the invention; use of such antibodies for drug targeting;
CC   and methods of screening for modulators of activity or expression of the
CC   polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC   antibodies and methods are useful for diagnosing, prognosing and treating
CC   cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC   atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC   neovascularistaion syndromes, scarring and uterine fibroids. They may
CC   also be useful in wound healing and in contraception. The present
CC   sequence represents a polypeptide of the invention.
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:               1.02e-134       Length:          253
Score:                   1364.00         Matches:         253
Percent Similarity:      100.00%         Conservative:    0
Best Local Similarity:   100.00%         Mismatches:      0
Query Match:             76.63%          Indels:          0
DB:                                      Gaps:            0

US-09-905-083A-30 (1-969)  x ADN39180 (1-253)
```

```
Qy   16   ATGGCAAGATCCCTTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT     75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1    MetAlaArgSerLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr     20

Qy   76   GCAGGAGAAGAAGCCAGGGTGACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCC    135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21   AlaGlyGluGluAlaArgValThrArgLeuLeuIleGlyProMetCysLysArgLeuSer  40

Qy   136  CACCCATGGCAGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGGCGGAGGCGTCCTGGTC   195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41   HisProTrpGlnTrpProLeuLeuSerGlyAsnGlnLeuHisTrpArgArgArgSerTrpSer  60

Qy   196  AATGAGGCTGGGTGCTCACTGCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG     255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61   AsnGluArgTrpValThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu     80

Qy   256  GGCAGTGATACGTGGGGCACAGGAGAGCTCAGAGGATCAAGGGCTCGAAGTCATTCCGC   315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81   GlySerAspThrTrpGlyThrGlyGluLeuArgGlySerLysAlaSerLysSerPheArg  100

Qy   316  CACCCGGCTACTCCACACAGACCATGTTAATGACCTCGTGAAGCTCAATAGC         375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101  HisProGlyTyrSerThrGlnThrMetLeuMetThrSerLysLeuAsnSer         120

Qy   376  CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGGCTGCCCTCCCGCTGCGAACCCCT  435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121  GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro  140

Qy   436  GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCAGAGCCCAGATGTGACCTTTCCC   495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141  GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro    160

Qy   496  TCTGACCTCATGTGCGTGGATGTCAAGTCTCATCTCCCCCAGAGACTGCACGAGAGGTTTAC  555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161  SerAspLeuMetCysValAspValLysSerHisLeuProGlnArgLeuHisGluArgPheTyr  180
```

```
Percent Similarity:      100.00%      Conservative:   0
Best Local Similarity:   100.00%      Mismatches:     0
Query Match:              76.63%      Indels:         0
DB:                            8      Gaps:           0

US-09-905-083A-30 (1-969) x ADL06515 (1-253)

Qy     16 ATGGGAAGATCCCTTCCTGCCCCTCGAGATCTTACTGCTGTCTATCCTTAGCCTTGGAAACT 75
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuLeuAlaLeuGluThr 20

Qy     76 GCAGGAGAAGAGCCCAGGTGACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCC 135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AlaGlyGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40

Qy    136 CACCCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCACTGCGGAGGGTCCTGGTC 195
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal 60

Qy    196 AATGAGCGGCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGATACACCGTGCACCTG 255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluThrValHisLeu 80

Qy    256 GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTGAAGTCATTCCGC 315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg 100

Qy    316 CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy    376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCCTCCCGCTGCGAACCCCCT 435
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy    436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCAGAGCCCAGATGTGACCTTTCCC 495
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160

Qy    496 TCTGACCTCATGTGCGTGAGTGCAAGTCAAGTCCCCCCAGGACTGCACGAGGTTTAC 555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 SerAspLeuMetCysValSerValAspValLysLeuIleSerProGlnAspCysThrArgValTyr 180

Qy    556 AAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAACGCC 615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200

Qy    616 TGCAATGGTGACTCAGGGGGACCGTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCC 675
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy    676 TGGGGAACTTCCCTTGCGCCAACCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy    736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATGC 774
          ||||||||||||||||||||||||||||||||||||||
Db    241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

RESULT 12
ADL06515
ID   ADL06515 standard; protein; 253 AA.
XX
AC   ADL06515;
XX
DT   20-MAY-2004 (first entry)
XX
DE   Human tumour-associated antigenic target (TAT) polypeptide #14.
XX
KW   Human; tumour-associated antigenic target; TAT; cell death; tumour;
KW   cancer; cytostatic.
XX
OS   Homo sapiens.
XX
PN   WO2004016225-A2.
XX
PD   26-FEB-2004.
XX
PF   19-AUG-2003; 2003WO-US025892.
XX
PR   19-AUG-2002; 2002US-0404809P.
PR   21-AUG-2002; 2002US-0405645P.
PR   23-SEP-2002; 2002US-0413192P.
PR   15-OCT-2002; 2002US-0419008P.
PR   15-NOV-2002; 2002US-0426847P.
PR   02-JUL-2003; 2003US-0484959P.
XX
PA   (GETH) GENENTECH INC.
XX
PI   Desauvage FJ, Frantz G, Hillan KJ, Polakis P, Polson A, Smith V;
PI   Spencer SD, Wu TD, Zhang Z;
XX
DR   WPI; 2004-257144/24.
DR   N-PSDB; ADL06435.
XX
PT   New antibody that binds to a tumor-associated antigenic target (TAT)
PT   polypeptide, useful for preparing a composition for diagnosing or
PT   treating cancer.
XX
PS   Claim 2; SEQ ID NO 95; 319pp; English.
XX
CC   The present invention relates to the isolation of human tumour-associated
CC   antigenic target (TAT) polynucleotide and polypeptide sequences. Also
CC   disclosed is an antibody that binds to a TAT polypeptide. The antibody is
CC   a monoclonal antibody, an antibody fragment, a chimeric antibody or a
CC   humanised antibody. It is conjugated to a growth inhibitory agent. It is
CC   produced in bacteria or in CHO cells and induces death of a cell to which
CC   it binds. The antibody is useful for preparing a composition for
CC   diagnosing or treating tumours and cancer. The present sequence
CC   represents a human TAT polypeptide of the invention.
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:   1.02e-134   Length: 253
Score: 1364.00   Matches: 253

RESULT 13
ADN04182
ID   ADN04182 standard; protein; 253 AA.
XX
AC   ADN04182;
XX
DT   01-JUL-2004 (first entry)
XX
DE   Antipsoriatic protein sequence #286.
XX
KW   antipsoriatic; gene therapy; psoriasis; diagnosis.
OS   Homo sapiens.
XX

```
PN   WO2004028479-A2.
XX
PD   08-APR-2004.
XX
PF   25-SEP-2003; 2003WO-US030907.
XX
PR   25-SEP-2002; 2002US-0414006P.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Bodary S, Clark H, Jackman J, Schoenfeld J, Williams PM, Wood WI;
PI   Wu TD;
XX
DR   WPI; 2004-305105/28.
DR   N-PSDB; ADN04181.
XX
PT   New PRO nucleic acid or polypeptide, useful for preparing a
PT   pharmaceutical composition for diagnosing or treating psoriasis in a
PT   mammal.
XX
PS   Claim 9; SEQ ID NO 576; 3069pp; English.
XX
CC   The invention relates to novel polynucleotide and polypeptides for
CC   treating psoriasis or a sequence having at least 80% identity to the
CC   above sequences. The nucleic acid is useful for preparing a composition
CC   for diagnosing or treating psoriasis in a mammal. This sequence
CC   corresponds to one of the polypeptides of the invention.
XX
SQ   Sequence 253 AA;

Alignment Scores:
Pred. No.:                1.02e-134      Length:          253
Score:                    1364.00        Matches:         253
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              76.63%         Indels:          0
DB:                       8              Gaps:            0

US-09-905-083A-30 (1-969) x ADN04182 (1-253)

Qy    16 ATGGCAAGATCCTTCTCCTGCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT     75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr  20

Qy    76 GCAGGAGAAGAAGCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAGAGAGGCTCC   135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgArgGlySer  40

Qy   136 CACCCCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC  195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal  60

Qy   196 AATGAGCGCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG  255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu  80

Qy   256 GGCAGTGATACGCTGGGGCGCACGAGGAGCTCAGGAGGATCAAGGCCTCGAAGTCATTCCGC  315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 GlySerAspThrLeuGlyArgThrArgSerSerGlyGlySerLysAlaSerLysSerPheArg 100

Qy   316 CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC  375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy   376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCCT  435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy   436 GGAACCACCTGTACTGTCTCCGCTGGGGCACTACCACGAGCCCAGGATGTGACCTTTCCC  495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160

Qy   496 TCTGACCTCACCGTGGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTAC  555
```

```
Db   161 SerAspLeuMetCysArgValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   556 AAGGACTTACTGCAAAATTCATCGTGCGCTGGCATCCCGACTCCAAGAAAAACGCC    615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200

Qy   616 TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCGTGGTGTCC  675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValLeuSer 220

Qy   676 TGGGGAACTTCCCTGCGGCCACCCAATGACCAGGAGTCTACACTGCAAGTGTGCAAG    735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221 TrpGlyThrPheProCysGlyGlnProAsnAspGlnGluSerThrLeuGlnValCysLys 240

Qy   736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC    774
          |||||||||||||||||||||||||||||||||||||||
Db   241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

```
RESULT 14
ADR72880
ID   ADR72880 standard; protein; 253 AA.
XX
AC   ADR72880;
XX
DT   02-DEC-2004  (first entry)
XX
DE   Human ovarian cancer-related tumour marker kallikrein 7 (hK7) protein.
XX
KW   kallikrein 7; tumour marker; ovarian cancer;
KW   epithelial ovarian carcinoma; human; serine protease; enzyme; KLK7.
XX
OS   Homo sapiens.
XX
PN   WO2004075713-A2.
XX
PD   10-SEP-2004.
XX
PF   26-FEB-2004; 2004WO-CA000281.
XX
PR   26-FEB-2003; 2003US-0450406P.
XX
PA   (MOUN ) MOUNT SINAI HOSPITAL.
XX
PI   Diamandis EP;
XX
DR   WPI; 2004-661815/64.
DR   N-PSDB; ADR72881, ADR72882.
XX
PT   Kallikrein markers detection method for detecting ovarian cancer in
PT   patient, involves detecting kallikrein markers and CA125 in sample
PT   obtained from patient, and comparing detected amounts with standard
PT   amounts.
XX
PS   Example 2; SEQ ID NO 10; 102pp; English.
XX
CC   The invention relates to a novel method for detecting a plurality of
CC   kallikrein markers associated with ovarian cancer. The method comprises
CC   obtaining a sample from a patient and detecting in the sample a plurality
CC   of kallikrein markers, and optionally carbohydrate antigen CA125, wherein
CC   the kallikrein markers are selected from the group consisting of
CC   kallikrein 5, kallikrein 6, kallikrein 7, kallikrein 8, kallikrein 10 and
CC   kallikrein 11. The detected amounts of the kallikrein markers are
CC   compared with standard amounts. The method of the invention may be useful
CC   for detecting kallikrein markers associated with ovarian cancer in a
CC   patient and thus for detecting ovarian cancer, particularly epithelial
CC   ovarian carcinoma. The current sequence is that of the human ovarian
CC   cancer-related tumour marker kallikrein 7 (hK7) protein of the invention
CC   which is a secreted serine protease and is encoded by DNA located at
CC   chromosome 19q13.4.
XX
SQ   Sequence 253 AA;
```

```
KW   immunomodulator; cytostatic; nootropic; neuroprotective;
KW   antiparkinsonian; antilipaemic; gene therapy; human disease;
KW   metabolic disorder; diabetes; obesity; infection; cachexia; cancer;
KW   neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;
KW   immune disorder; haematopoietic disorder; dyslipidaemia.
XX
OS   Homo sapiens.
XX
PN   WO200302942424-A2.
XX
PD   10-APR-2003.
XX
PF   02-OCT-2002; 2002WO-US031373.
XX
PR   02-OCT-2001; 2001US-0326483P.
PR   05-OCT-2001; 2001US-0327435P.
PR   05-OCT-2001; 2001US-0327449P.
PR   09-OCT-2001; 2001US-0327917P.
PR   09-OCT-2001; 2001US-0328029P.
PR   09-OCT-2001; 2001US-0328044P.
PR   09-OCT-2001; 2001US-0328056P.
PR   12-OCT-2001; 2001US-0328849P.
PR   15-OCT-2001; 2001US-0329414P.
PR   17-OCT-2001; 2001US-0301142P.
PR   18-OCT-2001; 2001US-0330309P.
PR   22-OCT-2001; 2001US-0341058P.
PR   24-OCT-2001; 2001US-0339266P.
PR   24-OCT-2001; 2001US-0343629P.
PR   29-OCT-2001; 2001US-0349575P.
PR   01-NOV-2001; 2001US-0346357P.
PR   17-APR-2002; 2002US-0373260P.
PR   19-APR-2002; 2002US-0373815P.
PR   19-APR-2002; 2002US-0373817P.
PR   19-APR-2002; 2002US-0373826P.
PR   19-APR-2002; 2002US-0373884P.
PR   22-APR-2002; 2002US-0374977P.
PR   16-MAY-2002; 2002US-0381037P.
PR   16-MAY-2002; 2002US-0381038P.
PR   16-MAY-2002; 2002US-0381042P.
PR   17-MAY-2002; 2002US-0381642P.
PR   28-MAY-2002; 2002US-0386056P.
PR   29-MAY-2002; 2002US-0383831P.
PR   25-JUN-2002; 2002US-0391335P.
PR   01-OCT-2002; 2002US-00262511.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Smithson G, Millet I, Peyman JA, Kekuda R, Ju J, Li L, Guo X;
PI   Patturajan M, Spytek KA, Edinger SR, Ellerman K, Malyankar UM;
PI   Ort T, Gorman L, Zerhusen BD, Anderson DW, Zhong M, Catterton E;
PI   Ji W, Miller CE, Rastelli L, Stone DJ, Pena CEA, Shenoy SG;
PI   Shimkets RA, Rothenberg ME, Leach MD, Agee ML, Berghs C, Dipippo VA;
PI   Eisen AJ, Gangolli EA, Rieger DK, Spaderna SK;
XX
DR   WPI; 2003-381626/36.
DR   N-PSDB; ADA05731.
XX
PT   New NOVX polypeptides and nucleic acids, useful for diagnosing,
PT   preventing or treating NOVX-associated disorders, e.g. diabetes, obesity,
PT   cancer or dyslipidemia, and in chromosome mapping, tissue typing or
PT   pharmacogenomics.
XX
PS   Claim 1; Page 169-170; 586pp; English.
XX
CC   The present invention describes NOVX proteins, where X can be 1 to 55
CC   (e.g. NOV1). Also described: (1) a composition comprising a polypeptide
CC   described above and a carrier; (2) a kit comprising, in one or more
CC   containers, the composition described above; (3) an isolated nucleic acid
CC   molecule which encodes a NOVX protein of the invention; (4) a vector
CC   comprising the nucleic acid molecule described above; (5) a cell
CC   comprising the above vector; (6) an antibody that immunospecifically
CC   binds to the polypeptide described above; (7) methods for determining the
CC   presence or amount of the above polypeptide or nucleic acid molecule in a
```

```
Alignment Scores:
Pred. No.:            1.02e-134        Length:           253
Score:               1364.00          Matches:          253
Percent Similarity:  100.00%          Conservative:     0
Best Local Similarity: 100.00%        Mismatches:       0
Query Match:         76.63%           Indels:           0
DB:                  8                Gaps:             0

US-09-905-083A-30 (1-969) x ADR72880 (1-253)

Qy     16 ATGGCAAGATCCCTTCTCCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetAlaArgSerLeuLeuLeuProLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy     76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGTCC 135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AlaGlyGluGluGluAlaGlnValThrArgLeuLeuMetAlaProHisValGlnGluValSer 40

Qy    136 CACCCATGGCAGGTGGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy    196 AATGAGCCGCTGGGTGTCCACTGCGCCCACTGCAAGATGAATGAGTACACCCGTGCACCTG 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AsnGluArgLeuGlyValHisCysAlaHisCysLysMetAsnGluTyrThrArgValHisLeu 80

Qy    256 GGGCAGTGATACGCTGGGGCACAGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 GlySerAspThrLeuGlyAlaGlnArgAlaGlnArgIleArgIleLysAlaSerLysSerPheArg 100

Qy    316 CACCCCGGCTACTCCACAGACCCATGTTAATGACCTCGTGAAGTCAATAGC 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuValLysSerLeuAsnSer 120

Qy    376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTGGCGAACCCCCT 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy    436 GGAACCACCTGCTACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160

Qy    496 TCTGACCTCATGTGCGTGGATGTCAAGTCATCTCCCCCCAGGACTGCACGAAGGTTTAC 555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 SerAspLeuMetCysValAspValLysSerLeuIleSerProGlnAspCysThrLysValTyr 180

Qy    556 AAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAAACGCC 615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProThrProArgLysLysAsnAla 200

Qy    616 TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAGAGGTCTGGTGTCC 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy    676 TGGGGAACTTCCCTTGCCGCCCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 TrpGlyThrPhePheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy    736 TTCACCAAGTGGATAAATGACACACATGAAAAAGCATCGC 774
          ||||||||||||||||||||||||||||||||||||||||
Db    241 PheThrLysTrpIleAsnAspThrMetLysHisArg 253

RESULT 15
ADA05732
ID   ADA05732 standard; protein; 250 AA.
XX
AC   ADA05732;
XX
DT   06-NOV-2003  (first entry)
XX
DE   Human NOV18a protein SEQ ID NO:92.
XX
KW   human; NOVX; antidiabetic; anorectic; antibacterial; virucide;
```

```
Qy  565 CTGAAAATTCCATGCTGCTGTGCGCTGGCATCCCCGACTCCAAGAAAAACGCCTGCAATGGT 624
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGly 200

Qy  625 GACTCAGGGGACCGTTGGTGTGCAGAGTACCCTGCAGAGGTCTGGTGTCCTGGGAACT 684
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201 AspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThr 220

Qy  685 TCCCTTGCGGCCACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTCACCAAG 744
            ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221 PheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLysPheThrLys 240

Qy  745 TGGATAAATGACACCATGAAAAAGCATCGC 774
            ||||||||||||||||||||||||||||||
Db  241 TrpIleAsnAspThrMetLysLysHisArg 250
```

Search completed: July 14, 2005, 17:25:05
Job time : 177.5 secs

CC sample; (8) methods for determining the presence of or predisposition to
CC a disease associated with altered levels of expression of the above
CC polypeptide or nucleic acid molecule in a first mammalian subject; (9) a
CC method of identifying an agent that binds to the polypeptide described
CC above; (10) a method for identifying a potential therapeutic agent for
CC use in treating a pathology that is related to an aberrant expression or
CC aberrant physiological interactions of the polypeptide; (11) a method of
CC screening for a modulator of activity or of latency or predisposition to
CC a pathology associated with the polypeptide; (12) a method for modulating
CC the activity of the polypeptide described above; (13) methods of treating
CC or preventing a pathology associated with the above polypeptide in a
CC mammal; and (14) a method for producing the above polypeptide. NOVX
CC sequences have antidiabetic, anorectic, antibacterial, virucide,
CC immunomodulator, cytostatic, nootropic, neuroprotective, antiparkinsonian
CC and antilipaemic activities, and can be used in gene therapy. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease. The polypeptide or the nucleic
CC acid molecule may be used to diagnose, treat or prevent metabolic
CC disorders such as diabetes or obesity, infections, cachexia, cancer,
CC neurodegenerative disorders such as Alzheimer's disease or Parkinson's
CC disease, immune disorders, haematopoietic disorders and various
CC dyslipidaemias. The nucleic acids can also be used as hybridisation
CC probes, in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. The present sequence represents a human NOVX from the
CC present invention.
XX
SQ Sequence 250 AA;

Alignment Scores:
Pred. No.:   Length: 250
Score: 3.07e-133   Matches: 250
Percent Similarity: 1350.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 75.84%   Gaps: 0
                      6

US-09-905-083A-30 (1-969) x ADA05732 (1-250)

```
Qy  25 TCCCTTCCTGCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGAGAA 84
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 SerLeuLeuLeuProLeuGlnIleLeuLeuLeuLeuSerLeuAlaLeuGluThrAlaGlyGlu 20

Qy  85 GAAGCCCAGGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCCACCATGG 144
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  21 GluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySerHisProTrp 40

Qy 145 CAGGTGGGCCCTGCTCAGTGCAATCAGTTCCACTGCGGAGGCGTCCTGGTCAATGAGCGC 204
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  41 GlnValAlaLeuLeuSerGlyAsnGlnAsnGlnPheHisCysGlyGlyValLeuValAsnGluArg 60

Qy 205 TGGGTGCTCACTGCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGAT 264
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  61 TrpValLeuThrAlaAlaHisCysLysLysMetAsnGluTyrThrValHisLeuGlySerAsp 80

Qy 265 ACGCTGGGCGACAGGAGGAGCTCAGAGAATCAAGGCCTCGAAGTCATTCCGACCCCCGGC 324
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  81 ThrLeuGlyAspArgArgArgAlaGlnArgIleLysAlaSerLysSerPheArgHisProGly 100

Qy 325 TACTCCACACAGACCATGTTAATGACCTCGTGGAAGCTCAATAGCCAGGCCAGG 384
           |||||||||||||||||||||||||||||||||||||||||||||||||||
Db 101 TyrSerThrGlnThrHisValAsnAspLeuMetLeuValLeuAsnSerGlnAlaArg 120

Qy 385 CTGTCATCCATGGTGAAGAAGAGTCAGGCTGCCCTCCGCTGCCGAACCCCTGGAACCACC 444
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 LeuSerSerMetValLysLysValArgGlnLeuProSerArgCysGluProProGlyThrThr 140

Qy 445 TGTACTGTCTCCGGCTGGGGCCACTACCACGAGCCCAGATGTGACCTTTCCCTGACCTC 504
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 141 CysThrValSerGlyTrpGlyThrTyrThrThrSerProAspValThrPheProSerAspLeu 160

Qy 505 ATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTACAAGGACTTA 564
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 161 MetCysValAspValLysLeuIleSerProArgThrAlaArgArgPheTyrLysValTyrLysAspLeu 180
```

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:   July 14, 2005, 17:16:29 ; Search time 37.5 Seconds
                              (without alignments)
                              3857.864 Million cell updates/sec

Title:          US-09-905-083A-30
Perfect score:  1780
Sequence:       1 ggattccgggctccatggc........aagaaacacaaaccctcag 969

Scoring table:  BLOSUM62
                Xgapop  10.0 , Xgapext  0.5
                Ygapop  10.0 , Ygapext  0.5
                Fgapop   6.0 , Fgapext  7.0
                Delop    6.0 , Delext   7.0

Searched:       513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters:    1027090

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09905083/runat_14072005_140354_10640/app_query.fasta_1.1159
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=Bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09905083 @CGN_1_1_46 @runat_14072005_140354_10640 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :   Issued_Patents_AA:*
        1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
        2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
        3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
        4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
        5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
        6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                            SUMMARIES
```

Result No.	Score	Query Match	Length	DB	Pred. No.	ID	Description
1	1371	77.0	265	4		US-09-949-016-7716	Sequence 7716, Ap
2	1364	76.6	253	2		US-08-557-146-2	Sequence 2, Appli
3	1364	76.6	253	2		US-08-824-874-3	Sequence 2, Appli
4	1364	76.6	253	3		US-09-154-344-2	Sequence 3, Appli
5	1364	76.6	253	3		US-08-930-188-2	Sequence 2, Appli
6	1364	76.6	253	3		US-09-210-084-3	Sequence 3, Appli
7	1364	76.6	253	5		US-09-764-762-3	Sequence 3, Appli
8	1364	76.6	253	2		PCT-US96-04294-2	Sequence 2, Appli
9	1235	69.4	225	2		US-09-027-337-4	Sequence 4, Appli
10	1235	69.4	225	4		US-09-644-600-4	Sequence 4, Appli
11	1235	69.4	225	4		US-09-654-600A-4	Sequence 4, Appli
12	1230	69.1	224	3		US-08-944-483-33	Sequence 33, Appl

13	1215.5	68.3	225	2	US-08-557-146-12	Sequence 12, Appl
14	1215.5	68.3	225	2	US-09-154-344-12	Sequence 12, Appl
15	829	46.6	154	3	US-09-261-416-7	Sequence 7, Appli
16	773	43.4	144	4	US-09-618-259-4	Sequence 4, Appli
17	614	34.5	260	3	US-09-025-059-3	Sequence 3, Appli
18	614	34.5	260	4	US-09-618-259-8	Sequence 8, Appli
19	612.5	34.4	270	4	US-09-949-016-7712	Sequence 7712, Ap
20	612.5	34.4	293	4	US-09-509-908-2	Sequence 2, Appli
21	582.5	32.7	260	3	US-09-07-526-2	Sequence 2, Appli
22	582.5	32.7	260	4	US-09-618-259-7	Sequence 7, Appli
23	582.5	32.7	290	4	US-09-949-016-8166	Sequence 8166, Ap
24	580.5	32.6	260	3	US-09-008-271A-7	Sequence 7, Appli
25	573.5	32.2	282	3	US-09-025-059-1	Sequence 1, Appli
26	571.5	32.1	250	4	US-09-205-258-427	Sequence 427, App
27	568.5	31.9	288	4	US-09-386-642-13	Sequence 13, Appl
28	567.5	31.9	254	4	US-09-439-313-525	Sequence 525, App
29	567.5	31.9	254	4	US-09-636-215-525	Sequence 525, App
30	567.5	31.9	254	4	US-09-685-166A-525	Sequence 525, App
31	567.5	31.9	254	4	US-09-679-426-525	Sequence 525, App
32	567.5	31.9	254	4	US-09-759-143-525	Sequence 525, App
33	567.5	31.9	254	4	US-09-651-236-525	Sequence 525, App
34	561	31.5	246	2	US-08-978-404B-44	Sequence 44, Appl
35	559.5	31.4	449	4	US-09-636-215-617	Sequence 617, App
36	559.5	31.4	449	4	US-09-685-166A-617	Sequence 617, App
37	559.5	31.4	449	4	US-09-679-426-617	Sequence 617, App
38	559.5	31.4	449	4	US-09-759-143-617	Sequence 617, App
39	559.5	31.4	449	4	US-09-651-236-617	Sequence 617, App
40	558	31.3	248	1	US-08-744-026-1	Sequence 1, Appli
41	558	31.3	248	2	US-09-102-732-1	Sequence 1, Appli
42	558	31.3	248	3	US-09-261-767-1	Sequence 1, Appli
43	558	31.3	268	3	US-08-824-874-1	Sequence 1, Appli
44	558	31.3	268	3	US-09-210-084-1	Sequence 1, Appli
45	558	31.3	268	4	US-09-764-762-1	Sequence 1, Appli

```
                            ALIGNMENTS

RESULT 1
US-09-949-016-7716
; Sequence 7716, Application US/0994916
; Patent No. 6812339
; GENERAL INFORMATION:
;  APPLICANT: VENTER, J. Craig et al.
;  TITLE OF INVENTION:  POLYMORPHISMS IN KNOWN GENES ASSOCIATED
;  TITLE OF INVENTION:  WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
;  FILE REFERENCE: CL001307
;  CURRENT APPLICATION NUMBER: US/09/949,016
;  CURRENT FILING DATE:  2000-04-14
;  PRIOR APPLICATION NUMBER: 60/241,755
;  PRIOR FILING DATE: 2000-10-20
;  PRIOR APPLICATION NUMBER: 60/237,768
;  PRIOR FILING DATE: 2000-10-03
;  PRIOR APPLICATION NUMBER: 60/231,498
;  PRIOR FILING DATE: 2000-09-08
;  NUMBER OF SEQ ID NOS: 207012
;  SOFTWARE: FastSEQ for Windows Version 4.0
;  SEQ ID NO 7716
;    LENGTH: 265
;    TYPE: PRT
;    ORGANISM: Human
US-09-949-016-7716

Alignment Scores:
Pred. No.:               9.84e-131        Length:         265
Score:                   1371.00          Matches:        254
Percent Similarity:      100.00%          Conservative:   1
Best Local Similarity:   99.61%           Mismatches:     0
Query Match:             77.02%           Indels:         0
DB:                      4                Gaps:           0

US-09-905-083A-30 (1-969) x US-09-949-016-7716 (1-265)

Qy    10 GGCTCCATGGCAAGATCCCTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTG 69
```

```
;  CURRENT APPLICATION DATA:
;    APPLICATION NUMBER: US/08/557,146
;    FILING DATE: 14-DEC-1995
;    CLASSIFICATION: 424
;  ATTORNEY/AGENT INFORMATION:
;    NAME: Sterner, Richard J.
;    REGISTRATION NUMBER: 35,372
;    REFERENCE/DOCKET NUMBER: 1103326-181
;  TELECOMMUNICATION INFORMATION:
;    TELEPHONE: (212) 819-8783
;    TELEFAX: (212) 354-8113
;  INFORMATION FOR SEQ ID NO: 2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH: 253 amino acids
;      TYPE: amino acid
;      TOPOLOGY: linear
;    MOLECULE TYPE: protein
US-08-557-146-2

Alignment Scores:
Pred. No.:                              Length:            253
Score:               4.97e-130          Matches:           253
Percent Similarity:  100.00%            Conservative:        0
Best Local Similarity: 100.00%          Mismatches:          0
Query Match:         76.63%             Indels:              0
DB:                                     Gaps:                2
```

US-09-905-083A-30 (1-969) x US-08-557-146-2 (1-253)

```
Qy   16 ATGGCAAGATCCCTTCCTCCGCCCTGCCAGATCTTACTGCTCTATCCTTAGCCTTGGAAACT   75
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuLeuSerLeuAlaLeuGluThr   20

Qy   76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCC   135
Db   21 AlaGlyGluGluAlaGlnGlyAspLeuLysIleIleAspGlyAlaProCysAlaArgGlySer   40

Qy  136 CACCCATGCAGGTGCCCTGCTCAGTGCAATCAGCTGCGGGAGGCGTCCTGGTC   195
Db   41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal   60

Qy  196 AATGAGCGGTGGTGCTCACTGCGCCCACTGCAAGATGAATGACAGTACACCGTG   255
Db   61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu   80

Qy  256 GGCAGTGATAGCGTGGGCGACAGGAGCTCAGAGATCAAGGCCTCAAGTCATTCCGC   315
Db   81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg   100

Qy  316 CACCCCGGCTACTCCACACAGCACCATGTTAATGACCTTAATGAGCTCGTGAAGCTCAATAGC   375
Db  101 HisProGlyTyrSerThrGlnHisHisValAsnAspLeuMetLeuValLysLeuAsnSer   120

Qy  376 CAGGCCAGGTGCTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCGCTGCGAACCCCCT   435
Db  121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysAlaGluPro   140

Qy  436 GGAACCACCTGTACTGTCTCCGGCTGGGCGCACTACCACCGAGCCCAGATGTGACCTTTCC   495
Db  141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro   160

Qy  496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCCCCCAGGACTGACGAAGGTTTAC   555
Db  161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr   180

Qy  556 AAGGACTTACTGAAAATTCCATGCTGTGCGCTGCATCCCGACTCCAAGAAAAACGCC   615
Db  181 LysAspLeuLeuGluAsnSerMetLeuCysValAlaGlyIlleProSerLysLysAsnAla   200

Qy  616 TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGTACCCTGCAAGGTCGTGGTGTCC   675
Db  201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer   220
```

```
Db   11 GlyThrMetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuLeuSerLeuAlaLeu   30

Qy   70 GAAACTGCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCATGTGCAAGA   129
Db   31 GluThrAlaGlyGluGluAlaGlnGlyAspLeuLysIleIleAspGlyAlaProCysAlaArg   50

Qy  130 GGCTCCCACCCATGCAGGTGCCCTGCTCAGTGCAATCAGCTGCGGGAGGCGTC   189
Db   51 GlySerHisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyVal   70

Qy  190 CTGTCAATGAGCGGTGGTGCTCACTGCGCCCACTGCAAGATGAATGACAGTACACCGTG   249
Db   71 LeuValAsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrVal   90

Qy  250 CACCTGGGCAGTGATAGCGTGGGCGACAGGAGCTCAGAGATCAAGGCCTCAAGTCA   309
Db   91 HisLeuGlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSer   110

Qy  310 TTCCGCCACCCCGGCTACTCCACACAGCACCATGTTAATGACCTTCGTGAAGCTC   369
Db  111 PheArgHisProGlyTyrSerThrGlnHisHisValAsnAspLeuMetLeuValLysLeu   130

Qy  370 AATAGCCAGGCCAGGTGCTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCGCTGCGAA   429
Db  131 AsnSerGlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGlu   150

Qy  430 CCCCTGGAACCACCTGTACTGTCTCCGGCTGGGCGCACTACCACCGAGCCCAGATGTGACC   489
Db  151 ProProGlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThr   170

Qy  490 TTTCCCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAG   549
Db  171 PheProSerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLys   190

Qy  550 GTTTACAAGGACTTACTGAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAA   609
Db  191 ValTyrLysAspLeuLeuGluAsnSerMetLeuCysValAlaGlyIlleProSerLysLys   210

Qy  610 AACGCCTGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTG   669
Db  211 AsnAlaCysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeu   230

Qy  670 GTGTCCTGGGAACTTCCCTTGCGGCCAACCCAATGACCCAGGAGTCTACACTCAAGTG   729
Db  231 ValSerTrpGlyThrPhePheProCysGlyGlnProAsnAspProGlyValTyrThrGlnVal   250

Qy  730 TGCAAGTTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC   774
Db  251 CysLysPheThrLysTrpIleIleAsnAspThrMetLysLysHisArg   265
```

```
RESULT 2
US-08-557-146-2
; Sequence 2, Application US/08557146
; Patent No. 5834290
; GENERAL INFORMATION:
;  APPLICANT: Egelrud, Torbjorn
;  APPLICANT: Hansson, Lennart
;  TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
;  TITLE OF INVENTION: Enzyme (SCCE)
;  NUMBER OF SEQUENCES: 17
;  CORRESPONDENCE ADDRESS:
;    ADDRESSEE: White & Case, Patent Department
;    STREET: 1155 Avenue of the Americas
;    CITY: New York
;    STATE: New York
;    COUNTRY: U.S.A.
;    ZIP: 10036-2787
;  COMPUTER READABLE FORM:
;    MEDIUM TYPE: Floppy disk
;    COMPUTER: IBM PC compatible
;    OPERATING SYSTEM: PC-DOS/MS-DOS
;    SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
```

Qy   1 TGGGGAACTTCCCTTGCGCGCCAACCCAATGACCCAGGAGTCTACACTGTGCAAG 735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLeuVal 240

Qy 736 TTCCACAGTGGATAAATGACACACATGAAAAAGCATCGC 774
          |||||||||||||||||||||||||||||||||||||||
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 3
US-08-824-874-3
; Sequence 3, Application US/08824874
; Patent No. 5962300
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Incyte Pharmaceuticals, Inc.
;   STREET: 3174 Porter Drive
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94304
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/824,874
;   FILING DATE: Filed Herewith
;   CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER:
;   FILING DATE:
; ATTORNEY/AGENT INFORMATION:
;   NAME: Billings, Lucy J.
;   REGISTRATION NUMBER: 36,749
;   REFERENCE/DOCKET NUMBER: PP-0252 US
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-855-0555
;   TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 253 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: GenBank
;   CLONE: 532504
US-08-824-874-3

Alignment Scores:
Pred. No.:   Length: 253
Score: 4.97e-130   Matches: 253
Percent Similarity: 1364.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 76.63%   Indels: 0
DB: 2   Gaps: 0

US-09-905-083A-30 (1-969) x US-08-824-874-3 (1-253)

Qy   16 ATGGCAAGATCCCCTTCCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MetAlaArgSerLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy   76 GCAGGAGAAGAAGCCCAGGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCC 135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AlaGlyGluGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40

Qy 136 CACCCATGCAGGTGGCCCTGCTCAGTGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195

RESULT 4
US-09-154-344-2
; Sequence 2, Application US/09154344
; Patent No. 5981256
; GENERAL INFORMATION:
; APPLICANT: Egelrud, Torbjorn
; APPLICANT: Hansson, Lennart
; TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
; TITLE OF INVENTION: Enzyme (SCCE)
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
;   ADDRESSEE: White & Case, Patent Department
;   STREET: 1155 Avenue of the Americas
;   CITY: New York
;   STATE: New York
;   COUNTRY: U.S.A.
;   ZIP: 10036-2787
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/154,344
;   FILING DATE: 16-SEP-1998
;   CLASSIFICATION:
; PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/557,146
;   FILING DATE: 14-DEC-1995
;   CLASSIFICATION:

Qy   1 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy   61 AATGAGCGCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80

Qy   81 GGCAGTGATACGCTGGGCACAGAGGAGCTCAGAGGATCAAGGCCTGAAGTCATTCCGC 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysSerLysSerPheArg 100

Qy 101 CACCCCGGCTACTCCACACAGACCATGTTAATGACCTCGTGAAGCTCAATAGC 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 101 HisProGlyTyrSerThrGlnThrMetHisValAsnAspLeuMetLysLeuAsnSer 120

Qy 121 CAGGCCAGCTGTCATCCATGGTGAAGAAGTCAGGCTGCCCTCCGCTGCGAACCCCT 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy 141 GGAACCACCTGTACTGTCTCCGGCTGGGGCCACTACCACGAGCCCAGATGTGACCTTTCCC 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160

Qy 161 TCTGACCTCATGTGCGTGATGTCAAGTCAATCTCCCCCAGGACTGCACGAAGGTTTAC 555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 161 SerAspLeuMetCysValValSerLeuIleSerProGlnAspCysThrLysValTyr 180

Qy 181 AAGGACTTACTGGAAAATTCCATGCTGTGCGCTGCATCCCGACTCCAAGAAAAACGCC 615
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProSerLysLysAsnAla 200

Qy 201 TGCAATGGTGACTCAGGGGACCGTTGGTGTCGAGAGGTACCCTGCAAGGTCTCGGTGTCC 675
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy 221 TGGGGAACTTCCCTTGCGCGCCAACCCAATGACCCAGGAGTCTACACTGTGCAAG 735
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy 241 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
          |||||||||||||||||||||||||||||||||||||||
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

ATTORNEY/AGENT INFORMATION:
  NAME: Sterner, Richard J.
  REGISTRATION NUMBER: 35,372
  REFERENCE/DOCKET NUMBER: 1103326-181
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (212) 819-8783
  TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 2:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 253 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
  MOLECULE TYPE: protein
US-09-154-344-2

Alignment Scores:
Pred. No.:   Length: 253
Score: 4.97e-130   Matches: 253
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.63%   Indels: 0
DB: 2   Gaps: 0

US-09-905-083A-30 (1-969) x US-09-154-344-2 (1-253)

Qy 16 ATGGCAAGATCCCTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
Qy 1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

RESULT 5
US-08-930-188-2
; Sequence 2, Application US/08930188
; Patent No. 6093397
GENERAL INFORMATION:
  APPLICANT: Dixon, Eric P.
  APPLICANT: Johnstone, Edward M.
  APPLICANT: Little, Sheila P.
  TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
  TITLE OF INVENTION: RELATED NUCLEIC ACIDS
  NUMBER OF SEQUENCES: 3
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Eli Lilly and Company
    STREET: Lilly Corporate Center
    CITY: Indianapolis
    STATE: Indiana
    COUNTRY: United States of America
    ZIP: 46285
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/930,188

Alignment Scores:
Pred. No.: Length: 253
Score: 4.97e-130 Matches: 253
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 76.63% Indels: 0
DB: 3 Gaps: 0

US-09-905-083A-30 (1-969) x US-08-930-188-2 (1-253)

```
;  TELEFAX:  415-845-4166
;  INFORMATION FOR SEQ ID NO: 3:
;  SEQUENCE CHARACTERISTICS:
;  LENGTH: 253 amino acids
;  TYPE: amino acid
;  STRANDEDNESS: single
;  TOPOLOGY: linear
;  IMMEDIATE SOURCE:
;  LIBRARY: GenBank
;  CLONE: 532504
US-09-210-084-3

Alignment Scores:
Pred. No.:              4.97e-130        Length:         253
Score:                  1364.00          Matches:        253
Percent Similarity:     100.00%          Conservative:   0
Best Local Similarity:  100.00%          Mismatches:     0
Query Match:            76.63%           Indels:         0
DB:                     3                Gaps:           0

US-09-905-083A-30 (1-969) x US-09-210-084-3 (1-253)

Qy    16 ATGGGAAGATCCCTTCTCCGCCCCTCGAGATCTTACTGTGTATCCTTAGCCTTGGAAACT 75
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy    76 GCAGGAGAAGAGCCCAGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCC 135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer 40

Qy   136 CACCCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy   196 AATGAGCGCTGGGTGCTCACTGCGCCCACTGCCAAGATGAATGAGTACACCGTGCACCTG 255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80

Qy   256 GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC 315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg 100

Qy   316 CACCCCGGCTACTCCACACAGACACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC 375
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy   376 CAGGCCAGGTGTCATCCATGGTGAAGAAAGTCAGGGTGCCCTGCCTCCCGGTGCGAACCCCT 435
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GlnAlaArgCysHisProTrpSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy   436 GGAACCACCGTACTGTCTCCGGCTGGGCGACTACCACGAGCCCAGATGTGACCTTTCCC 495
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 GlyThrThrValLeuSerArgLeuGlyTrpGlyTrpThrThrSerProAspValThrPhePro 160

Qy   496 TCTGACCTCATGTGCGTGATGTCAAGCTCATCTCCCCCAGGACTGACGAAGGTTAC 555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 SerAspLeuMetCysValAspSerValLysSerIleSerProGlnAspCysThrLysValTyr 180

Qy   556 AAGGACTTACTGGAAAATTCCATGCTGCTCGTGGCATCCCGACTCCAAGAAAACGCC 615
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LysAspLeuLeuGluAsnSerMetLeuCysValAlaGlyIleProAspSerLysAsnAla 200

Qy   616 TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGTCTGGTGTCC 675
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy   676 TGGGGAACTTCCCTTGCGGCCAACCCAATGACCCCAGAGTCTACACTCCAAGTGTGCAAG 735
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy   736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
          |||||||||||||||||||||||||||||||||||||||
Db   241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

RESULT 6
US-09-210-084-3
; Sequence 3, Application US/09210084
; Patent No. 6197511
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,084
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/824,874
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555

RESULT 7
US-09-764-762-3
; Sequence 3, Application US/09764762
; Patent No. 6472195
; GENERAL INFORMATION:
;   APPLICANT: Hillman, Jennifer L.
;   APPLICANT: Lal, Preeti
;   TITLE OF INVENTION: NOVEL KALLIKREIN
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Incyte Pharmaceuticals, Inc.
;   STREET: 3174 Porter Drive
;   CITY: Palo Alto
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94304
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/764,762
;   FILING DATE: 16-Jan-2001
;   CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 09/210,084
;   FILING DATE: <Unknown>
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Billings, Lucy J.
;   REGISTRATION NUMBER: 36,749
;   REFERENCE/DOCKET NUMBER: PF-0252 US
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-855-0555
;   TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 253 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   IMMEDIATE SOURCE:
;   LIBRARY: GenBank
;   CLONE: 532504
;   SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-764-762-3

Alignment Scores:
  Pred. No.:   4.97e-130   Length: 253
  Score: 1364.00   Matches: 253
  Percent Similarity: 100.00%   Conservative: 0
  Best Local Similarity: 100.00%   Mismatches: 0
  Query Match: 76.63%   Indels: 0
  DB: 4   Gaps: 0

US-09-905-083A-30 (1-969) x US-09-764-762-3 (1-253)

Qy   16 ATGGCAAGATCCCTTCTCCTGCCCCTGCCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MetAlaArgSerLeuLeuLeuProLeuProAspLeuThrAlaIleLeuSerLeuAlaLeuGluThr 20

Qy   76 GCAGGAGAAGCCAGGTGACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCC 135
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AlaGlyGluGlyGlnAlaGlnIleIleIleAspGlyAlaProAlaProCysAlaArgGlySer 40

Qy 136 CACCCATGCAGGTGGCCCTGCTCAGTGGCAATCAGCTTCCACTGCGGAGGCGTCCTGGTC 195
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 HisProTrpGlnAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy 196 AATGAGCCGCTGGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG 255
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80

Qy 256 GGCAGTGATACGCTGGCCGACAGGAGAGAGCTCAGAGGATCAAGGCCTGAGTCATTCCGC 315
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 GlySerAspThrLeuGlyAspArgArgGluLeuArgGlySerLysAlaSerLysSerPheArg 100

Qy 316 CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCGTGAAGCTCAATAGC 375
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuValLysLeuAsnSer 120

Qy 376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGGCCTGCCTCCCGCTGCGAACCCCT 435
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysSerGluProPro 140

Qy 436 GGAACCACCTGTACTGTCTCCGGCTGGGCACTACCACGAGCCCAGATGTGACCTTTCCC 495
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160

Qy 496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTCGCACGAAGGTTTAC 555
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180

Qy 556 AAGGACTTACTGGAAAATTCCATGCTGCGCTGGCATCCCGACTCCAAGAAAAAACGCGC 615
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200

Qy 616 TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCC 675
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy 676 TGGGGAACTTTCCCTTGGCGGCCAACCCATGACCCAGGAGTGTCTACTCAAGTGTGCAAG 735
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy 736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
         |||||||||||||||||||||||||||||||||||||||
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 8
PCT-US96-04294-2
; Sequence 2, Application PC/TUS9604294
; GENERAL INFORMATION:
;   APPLICANT: Dixon, Eric P.
;   APPLICANT: Johnstone, Edward M.
;   APPLICANT: Little, Sheila P.
;   TITLE OF INVENTION: AMYLOID PRECURSOR PROTEIN PROTEASE AND
;   TITLE OF INVENTION: RELATED NUCLEIC ACIDS
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Eli Lilly and Company
;   STREET: Lilly Corporate Center
;   CITY: Indianapolis
;   STATE: Indiana
;   COUNTRY: United States of America
;   ZIP: 46285
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: PCT/US96/04294
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/416,257
;   FILING DATE: 04-APR-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Blalock, Donna K.
;   REGISTRATION NUMBER: 38,082
;   REFERENCE/DOCKET NUMBER: X9239
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 317-277-1090
;   TELEFAX: 317-276-3861

```
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 253 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
PCT-US96-04294-2

Alignment Scores:
Pred. No.:              4.97e-130      Length:        253
Score:                  1364.00        Matches:       253
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            76.63%         Indels:        0
DB:                                    Gaps:          5

US-09-905-083A-30 (1-969) x PCT-US96-04294-2 (1-253)

Qy    16 ATGGCAAGATCCCTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCCTTGGAAACT    75
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr    20

Qy    76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCC   135
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer    40

Qy   136 CACCCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGCTTCACTGCGGAGGCGTCCTGGTC   195
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal    60

Qy   196 AATGAGCGGCTGGGTGCTCACTGCGCCCACTGACAAGATGAATGAGTACACCGTGCACCTG   255
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 AsnGluArgTrpValLeuThrAlaAlaHisCysAlaHisCysAsnGluTyrThrValHisLeu    80

Qy   256 GGCAGTGATACGCTGGGCCACAGAGAGCTCAGAGGCTCAAGGCCTCGAAGTCATTCCGC   315
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 GlySerAspThrLeuGlyHisArgGluArgGlnArgLeuLysLysSerGlnGlySerPheArg   100

Qy   316 CACCCCGGCTACTCCACAGACCCATGTAATGACCTCGTGAAGCTCAATAGC   375
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer   120

Qy   376 CAGGCCAGGTCGTCATCCATGGTGAAGAAAGTCAGGCTGCCCCTGCGACCCCCT   435
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro   140

Qy   436 GGAACCACCTGTACTGTCTCCGGCTGGGACTACCGTGGGCCAGATGTGACCTTTCCC   495
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 GlyThrThrCysThrValSerGlyTrpGlyThrThrTrpSerProAspValThrPhePro   160

Qy   496 TCTGACCTTCATGTGCGTGATGTCAAGCTCATCTCCCCAGGACTGCACGAAGGTTTAC   555
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr   180

Qy   556 AAGGACTTACTGGAAAATTCCATGTGCGCTGGCATCCCCAGAGAAAAACGCC   615
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LysAspLeuLeuGluAsnSerMetCysAlaGlyIleProArgSerLysLysAsnAla   200

Qy   616 TGCAATGGTGACTCAGGGGACCGGTTGGTGCAGAGGTACCTGCAAGTCTGGTGTCC   675
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer   220

Qy   676 TGGGGAACTTCCCTGCGCCACCCAATGACCCCAGAGTCTTACACTCAAGTGTGCAAG   735
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   221 TrpGlyThrPhePheProGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys   240

Qy   736 TTCACCAAGTGGATAAATGACACCATGAAAAAACATCGC   774
         |||||||||||||||||||||||||||||||||||||||
Db   241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg   253

RESULT 9
US-09-027-337-4
```

```
; Sequence 4, Application US/09027337B
; Patent No. 5972616
; GENERAL INFORMATION:
;   APPLICANT: O'Brien, Timothy J.
;   APPLICANT: Tanimoto, Hirotoshi
;   TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease Overexpressed in
;   TITLE OF INVENTION: Breast and Ovarian Carcinomas
;   FILE REFERENCE: D6064
; CURRENT APPLICATION NUMBER: US/09/027,337B
; CURRENT FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 13
; SEQ ID NO 4
;   LENGTH: 225
;   TYPE: PRT
;   ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Serine protease catalytic domain of Scce homologous to
; OTHER INFORMATION: similar domain in TADG-15
US-09-027-337-4

Alignment Scores:
Pred. No.:              6.11e-117      Length:        225
Score:                  1235.00        Matches:       225
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            69.38%         Indels:        0
DB:                                    Gaps:          2

US-09-905-083A-30 (1-969) x US-09-027-337-4 (1-225)

Qy   100 AAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCCCACCCATGGCAGGTGGCCCTGCTC   159
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 LysIleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGlnValAlaLeuLeu    20

Qy   160 AGTGGCAATCAGCTTCACTGCGGAGGCGTCCTGGTCAATGAGCGGCTGGGTGCTCACTGCC   219
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 SerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAla    40

Qy   220 GCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACGCTGGGCGACAGG   279
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 AlaHisCysLysMetAsnGluTyrThrValHisLeuGlySerAspThrLeuGlyAspArg    60

Qy   280 AGAGCTCAGAGGATCAAGGCCTGAAGTCATTCCGCCACCCCGGCTACTCCACACAGACC   339
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ArgAlaGlnArgIleLysAlaSerIleSerPheArgHisProGlyTyrSerThrGlnThr    80

Qy   340 CATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGGCCAGGCTGTCATCCATGGTG   399
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 HisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerSerMetVal   100

Qy   400 AAGAAAGTCAGGCTGCCTCGCTGGAACCCCTGGAACCACCTGTACTGTCTCCGGC   459
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 LysLysValArgLeuProSerArgCysGluProGlyThrThrCysThrValSerGly   120

Qy   460 TGGGGCACTACCAGAGCCAGATGTGACCTTTCCCTGACCTTCATGTGCGTGATGTC   519
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TrpGlyThrThrSerProAspValThrPheProSerAspLeuMetCysValAspVal   140

Qy   520 AAGCTCATCTCCCCAGGACTGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATG   579
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 LysLeuIleSerProGlnAspCysThrLysValTyrLysAspLeuLeuGluAsnSerMet   160

Qy   580 CTGTGCGCTGGCATCCCCGACTCCAAGAAAAAACGCCTGCAATGGTGACTCAGGGGACCG   639
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 LeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyPro   180

Qy   640 TTGGTGTGCAGAGGTACCCTGCAAGTGTGTCCTGGTGTCCTGGGGAACTTCCCTGCGGCCAA   699
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   181 LeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPhePheProCysGlyGln   200

Qy   700 CCCAATGACCCCAGAGTCTTACACTCAAGTGTGCAAGTTCACCAAGTGGATAAATGACACC   759
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   201 ProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThr   220
```

Qy   760 ATGAAAAAGCATGC 774
          |||||||||||||
Db   221 MetLysLysHisArg 225

RESULT 10
US-09-644-600-4
; Sequence 4, Application US/09644600
; Patent No. 6451500
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotoshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/644,600
; CURRENT FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: 09/421,213
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: 09/027,337
; PRIOR FILING DATE: 1998-02-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-644-600-4

Alignment Scores:
Pred. No.:   Length: 225
Score: 6.11e-117   Matches: 225
Percent Similarity: 1235.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 69.38%   Gaps: 0
                        4

US-09-905-083A-30 (1-969) x US-09-644-600-4 (1-225)

Qy   100 AAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCCCACCCATGGCAGGTGGCCCTGCTC 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 LysIleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGlnValAlaLeuLeu 20

Qy   160 AGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTCAATGAGCGCTGGTGCTCACTGCC 219
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 SerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAla 40

Qy   220 GCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACGCTGGGCGACAGG 279
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 AlaHisCysLysMetAsnGluTyrThrValHisLeuGlySerAspThrLeuGlyAspArg 60

Qy   280 AGAGCTCAGAGGATCAAGGCCTGAAGTCATTCCGCCACCACCCGGCTACTCCACACAGACC 339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ArgAlaGlnArgIleLysAlaSerLysSerPheArgHisHisProGlyTyrSerThrGlnThr 80

Qy   340 CATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGGCCGTGCTCATCCATGGTG 399
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 HisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerSerMetVal 100

Qy   400 AAGAAAGTCAGGCTGCCCTCCCGCTGGAACCCCCTGGAACCACTGTGTCTCCCGGC 459
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 LysLysValArgLeuProSerArgCysGluProProGlyThrThrCysThrValSerGly 120

Qy   460 TGGGGCACTACCACGAGCCCAGATGTGACCTTTCCTCTGACCTCATGTGCGTGGATGTC 519
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TrpGlyThrThrSerProAspValThrPheProSerAspLeuMetCysValAspVal 140

Qy   520 AAGCTCATCTCCCCAGGACTGCACGGAGGTTTACAAGGACTTACTGGAAAATTCCATG 579
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 LysLeuIleSerProGlnAspCysThrGluValTyrLysAspLeuGlyAsnSerMet 160

Qy   580 CTGTGCGCTGGCATCCCGACTCCCAGAGAAAAACGCCTCAAGATGGTGTGACTCAGGGGGACCG 639
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 LeuCysAlaGlyIleProThrProSerLysAsnAlaCysAsnGlyAspSerGlyGlyPro 180

Db   161 LeuCysAlaGlyIleProThrProSerLysAsnAlaCysAsnGlyAspSerGlyGlyPro 180
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   640 TTGGTGTGCAGAGGTACCCGCAAGCTCGGTGTCTGGTCTGGGGAACTTCCCTTGCCGGCCAA 699

Db   181 LeuValCysArgGlyThrArgGlyLeuValSerTrpGlyLeuLeuValSerTrpGlyGln 200
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy   700 CCCAATGACCCAGGAGTCTACACTCAAGTGCGAAGTTCACCAGTGGATAAATGACACC 759

Db   201 ProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysPheThrLysTrpIleAsnAspThr 220

Qy   760 ATGAAAAAGCATCGC 774
          |||||||||||||
Db   221 MetLysLysHisArg 225

RESULT 11
US-09-654-600A-4
; Sequence 4, Application US/09654600A
; Patent No. 6649741
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Tanimoto, Hirotoshi
; TITLE OF INVENTION: TADG-15: An Extracellular Serine Protease
; TITLE OF INVENTION: Overexpressed in Carcinomas
; FILE REFERENCE: D6064CIP/D
; CURRENT APPLICATION NUMBER: US/09/654,600A
; CURRENT FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/421,213
; 09/027,337
; 1998-02-20
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 225
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: SCCE
US-09-654-600A-4

Alignment Scores:
Pred. No.:   Length: 225
Score: 6.11e-117   Matches: 225
Percent Similarity: 1235.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 69.38%   Gaps: 0
                        4

US-09-905-083A-30 (1-969) x US-09-654-600A-4 (1-225)

Qy   100 AAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCCCACCCATGGCAGGTGGCCCTGCTC 159
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 LysIleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGlnValAlaLeuLeu 20

Qy   160 AGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTCAATGAGCGCTGGTGCTCACTGCC 219
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 SerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAla 40

Qy   220 GCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACGCTGGGCGACAGG 279
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 AlaHisCysLysMetAsnGluTyrThrValHisLeuGlySerAspThrLeuGlyAspArg 60

Qy   280 AGAGCTCAGAGGATCAAGGCCTGAAGTCATTCCGCCACCACCCGGCTACTCCACACAGACC 339
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 ArgAlaGlnArgIleLysAlaSerLysSerPheArgHisHisProGlyTyrSerThrGlnThr 80

Qy   340 CATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGGCCGTGCTCATCCATGGTG 399
          |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 HisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerSerMetVal 100

Qy   400 AAGAAAGTCAGGCTGCCCTCCCGCTGGAACCCCCTGGAACCACTGTGTCTCCGGC 459
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 LysLysValArgLeuProSerArgCysGluProProGlyThrThrCysThrValSerGly 120

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: No. 6232456e
US-08-944-483-33

Alignment Scores:
Pred. No.:                          Length:        224
Score:                 1.96e-116    Matches:       224
Percent Similarity:    1230.00      Conservative:    0
Best Local Similarity: 100.00%      Mismatches:      0
Query Match:           100.00%      Indels:          0
DB:                    69.10%       Gaps:            0
                        3

US-09-905-083A-30 (1-969) x US-08-944-483-33 (1-224)

Qy  103  ATTATTGATTGGCGCCCATGTGCAAGAGAGGCTCCCACCCATGGAGCAGGTGGCCCTGCTCAGT  162
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  IleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGlnValAlaLeuLeuSer   20

Qy  163  GGCAATCAGCTCCACTGCCGGAGGGTCCTGGTCAATGAGCGCCTGGTGCTCCACTGCCGCC  222
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21  GlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAlaAla   40

Qy  223  CACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGAGCGATACCGCTGGGCGACAGGAGA  282
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41  HisCysLysMetAsnGluTyrThrValHisLeuGlySerAspThrLeuGlyGlyAspArgArg   60

Qy  283  GCTCAGAGGATCAAGGCCTCGAAGTCATTCCGCACCCCGGCTACTCCACACAGACCCAT  342
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AlaGlnArgIleLysSerLysSerPheArgThrProAlaThrProHisGlnThrHis   80

Qy  343  GTTAATGACCTCTGTGAAGCTCAATAGCCAGGCCAGGCTGTCATCCATGTGAAG  402
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81  ValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerSerMetValLys  100

Qy  403  AAAGTCAGGCTGCCCTCCCGCTGCGAACCCCTGGAACCACCTGTACTGTCTCCGGCTGG  462
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101  LysValArgLeuProSerArgCysGluProProGlyThrThrCysTyrCysLeuArgLeuSerGlyTrp  120

Qy  463  GGCACTACCACGAGCCCAGATGTGACCTTTCCCTGACCTCATGTGCGTGGATGTCAAG  522
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GlyThrThrThrSerProAspValThrPheProSerAspLeuMetCysValAspValLys  140

Qy  523  CTCATCTCCCCCAGGACTGCACGAAGGTTACAGGACTACTGGAAAATTCCATGTCTG  582
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141  LeuIleSerProGlnAspCysThrLysValTyrLysValTyrLysAspLeuLeuGluAsnSerMetLeu  160

Qy  583  TGCGCTGGCATCCCGACTCCAAGAAAACGCCTGCAATGGTGACTCAGGGGGACCGTTG  642
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161  CysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyProLeu  180

Qy  643  GTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCCTGGGGAACTTTCCCTGCGGCCAACCC  702
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  ValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPheProCysGlyGlnPro  200

Qy  703  AATGACCACCAGGAGTCTACACTCCAAGTGTCCAAGTTCACCAAGTTGGATAAATGACACCATG  762
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201  AsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThrMet  220

Qy  763  AAAAAGCATCGC  774
              ||||||||||||
Db  221  LysLysHisArg  224
```

```
Qy  460  TGGGGCACTACCACGAGCCCAGATGTGACCTTTCCCTGACCTCATGTGCGTGGATGTC  519
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  TrpGlyThrThrThrSerProAspValThrPheProSerAspLeuMetCysValAspVal  140

Qy  520  AAGCTCATCTCCCCCAGGACTGCACGAAGGTTACAGGACTACTGGAAAATTCCATG  579
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141  LysLeuIleSerProGlnAspCysThrLysValTyrLysValTyrLysAspLeuLeuAsnSerMet  160

Qy  580  CTGTGCGCTGGCATCCCGACTCCCAAGAAAAACGCCTGCAATGGTGACTCAGGGGGACCG  639
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161  LeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyPro  180

Qy  640  TTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCCTGGGGAACTTTCCCTGCGGCCAA  699
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPheProCysGlyGln  200

Qy  700  CCCAATGACCCCAGGAGTCTACACTCCAAGTGTGCAAGTTCACCAAGTGGATAAATGACACC  759
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201  ProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThr  220

Qy  760  ATGAAAAAGCATCGC  774
              |||||||||||||||
Db  221  MetLysLysHisArg  225
```

CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case, Patent Department
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 225 amino acids
TYPE: amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
US-08-557-146-12

Alignment Scores:

Pred. No.:	5.85e-115	Length:	225
Score:	1215.50	Matches:	223
Percent Similarity:	99.56%	Conservative:	1
Best Local Similarity:	99.11%	Mismatches:	0
Query Match:	68.29%	Indels:	1
DB:		Gaps:	1

US-09-905-083A-30 (1-969) x US-08-557-146-12 (1-225)

```
Qy 103 ATTATTGATGGGCGCCCATGTGCAAGAGGCTCCACCATGGCAGGTGGCCCTGCTCAGT 162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 IleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGlnValAlaLeuLeuSer  20

Qy 163 GGCAATCAGCTC---CACTGCGGAGGCGTCCTGGTCAATGAGCGCTGGGTGCTCACTGCC 219
          ||||||||||    ||||||||||||||||||||||||||||||||||||||
Db  21 GlyAsnGlnLeuHisHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAla  40

Qy 220 GCCCACTGCAAGATGAATGAAGTAGTACACCGTGCACCTGGCGAGTGCAGTGATACGCTGGGCGACAGG 279
          ||||||||||||||||||||||||||||||||||||||||||||||||
Db  41 AlaHisCysLysMetAsnGluValTyrThrValHisLeuGlySerAspThrLeuGlyAspArg  60

Qy 280 AGAGCTCAGAGATCAAGGCTTCGAAGTCATTCCGCAACCACCCCGGCTACTCCACACAGACC 339
          ||||||||||||||||||||||||||||||||||||||||||||
Db  61 ArgAlaGlnArgIleLysSerLysSerPheArgHisProGlyTyrSerThrGlnThr  80

Qy 340 CATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGGCCAGGCTGTCATCCATGGTG 399
          ||||||||||||||||||||||||||||||||||||||||||||
Db  81 HisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerSerMetVal 100

Qy 400 AAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCCTGGAACCACCTGTACTGTCTCCGGC 459
          ||||||||||||||||||||||||||||||||||||||||||||
Db 101 LysLysValArgLeuProSerArgCysGluProProGlyThrThrCysThrValSerGly 120

Qy 460 TGGGGCACTACCACGAGCCCAGATGTGACCTTTCCCTCTGACCTCATGTGCGTGGATGTC 519
          ||||||||||||||||||||||||||||||||||||||||||||
Db 121 TrpGlyThrThrSerProAspValThrProSerAspThrLeuMetCysValAspVal 140

Qy 520 AAGCTCATCTCCCCCAGGACTGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATG 579
          ||||||||||||||||||||||  ::||||||||||||||||||
Db 141 LysLeuIleSerProGlnAspCysThrGluValTyrLysTyrGluLeuLeuAsnSerMet 160
```

```
Qy 580 CTGTGCGGCTGGCATCCCCGACTCCAAGAGAAAACGCCTGCAATGTGACTCAGGGGGACCG 639
          |||||||||||||||||||||||||||||||||||||||||||
Db 161 LeuCysAlaGlyIleLeuProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyPro 180

Qy 640 TTGGTGTGCAGAGTACCCTGCCAAGGTCTGGTGTCTGGGAACTTTCCCTTGGGCCAA 699
          ||||||||||||||||||||||||||||||||
Db 181 LeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyLeuValSerProCysGlyGln 200

Qy 700 CCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTGCGATAAATGACACC 759
          ||||||||||||||||||||||
Db 201 ProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThr 220

Qy 760 ATGAAAAAGCATCGC 774
          ||||||||||||
Db 221 MetLysLysHisArg 225
```

RESULT 14
US-09-154-344-12
; Sequence 12, Application US/09154344
; Patent No. 5981256
GENERAL INFORMATION:
APPLICANT: Egelrud, Torbjorn
APPLICANT: Hansson, Lennart
TITLE OF INVENTION: Recombinant Stratum Corneum Chymotryptic
TITLE OF INVENTION: Enzyme (SCCE)
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: White & Case, Patent Department
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2787
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/154,344
FILING DATE: 16-SEP-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/557,146
FILING DATE: 14-DEC-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sterner, Richard J.
REGISTRATION NUMBER: 35,372
REFERENCE/DOCKET NUMBER: 1103326-181
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 819-8783
TELEFAX: (212) 354-8113
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 225 amino acids
TYPE: amino acids
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: polypeptide
US-09-154-344-12

Alignment Scores:

Pred. No.:	5.85e-115	Length:	225
Score:	1215.50	Matches:	223
Percent Similarity:	99.56%	Conservative:	1
Best Local Similarity:	99.11%	Mismatches:	0
Query Match:	68.29%	Indels:	1
DB:		Gaps:	1

US-09-905-083A-30 (1-969) x US-09-154-344-12 (1-225)

```
Score:                  829.00          Matches:        154
Percent Similarity:     100.00%         Conservative:     0
Best Local Similarity:  100.00%         Mismatches:       0
Query Match:            46.57%          Indels:           0
DB:                      3              Gaps:             0

US-09-905-083a-30 (1-969) x US-09-261-416-7 (1-154)
```

```
Qy   103 ATTATTGATGGCGGCGCCCTCCCCATGGCAGAGGCTCCCACCCCAGAGGCTGGCCCTGCTCAGT 162
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 IleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGlnValAlaLeuLeuSer 20

Qy   163 GGCAATCAGTC---CACTGCGGAGGCGTCCTGGTCAATGAGCGCTGGTGCTCACTGCC 219
             |||||||||     |||||||||||||||||||||||||||||||||||
Db    21 GlyAsnGlnLeuHisHisCysGlyGlyValLeuValAsnGluArgTrpValLeuThrAla 40

Qy   220 GCCCACTGCCAAGATGAATGAGTACACCGTGCACCTGGCAGTGATACGCTGGGCGACAGG 279
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 AlaHisCysLysMetAsnGluTyrThrValHisLeuGlySerAspThrGlyGlyAspArg 60

Qy   280 AGAGCTCAGGAGGATCAAGGCCTGGAAGTCAATTCCGCCACCCCGGCTACTCCACACAGACC 339
             ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 ArgAlaGlnArgIleLysAlaSerLysLysSerPheArgHisProGlyTyrSerThrGlnThr 80

Qy   340 CATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGGCTGTCATCCATGGTG 399
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 HisValAsnAspAspLeuMetLeuValLeuLysLeuAsnSerGlnAlaArgLeuSerSerMetVal 100

Qy   400 AAGAAAGTCAGGTGCGCCTCCCGCTGCGAACCCCTGAAACCACCTGTACTGTCTCCCGC 459
             |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 LysLysValValArgLeuProSerArgCysGluProProGlyThrThrCysThrValSerGly 120

Qy   460 TGGGGCACTACCACGAGCCCAGATGTGACCTTTCCCTGACCTCATGTGCGTGGATGTC 519
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 TrpGlyThrThrSerProAspValThrPheProSerAspLeuMetCysValAspValAspVal 140

Qy   520 AAGCTCATCTCCCCCAGACTGCACGAGGTTTACAAGGACTTACTGGAAAATTCCATG 579
             |||:||||||||||||||||||||||||||||||||
Db   141 LysLeuIleSerProGlnAspCysTrpGlnValThrTyrLysGluValAspAspCysTrpProGlyGlySerMet 160

Qy   580 CTGTGCGCTGGCATCCCGACTCCCAAGAAGAAAACGCCTGCAATGGTGACTCAGGGGACCG 639
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 LeuCysAlaGlyIleProAspSerLysLysProAspSerLysLysThrGluValTyrLysAspCysTrpProCysGlyGlyPro 180

Qy   640 TTGGTGTGCAGAGTACCCTGCAAGGTCTGGTGTCCCTTGCGGCCAA 699
             |||||||||||||||||||||||||||||||||||||||||
Db   181 LeuValCysAlaGlyThrLeuGlnGlyLeuLeuValSerTrpGlyLeuLeuValSerTrpGlyThrTrpPhePheProCysGlyGln 200

Qy   700 CCCAATGACCCAGGAGTCTACACCCAAGTTGCAAGTCCCAGTGGATAAATGACACC 759
             |||||||||||||||||||||||||||||||||||||||||||||||
Db   201 ProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThr 220

Qy   760 ATGAAAAAGCATCGC 774
             |||||||||||
Db   221 MetLysLysHisArg 225
```

```
Qy   178 TGCGGAGGCGTCCTGGTCAATGAGCGCTGGTGCTCACTGCGGCCCACTGCCAAGATGAAT 237
             ||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 CysGlyGlyValLeuValAsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsn 20

Qy   238 GAGTACACCGTGCACCTGGCAGTGATACGCTGGGCGACAGGAGAGCTCAGGAGGATCAAG 297
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 GluTyrThrValHisLeuGlySerAspThrGlyGlyAspArgArgAlaGlnArgIleLys 40

Qy   298 GCCTGGAAGTCAATTCCGCCACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCATG 357
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 AlaSerLysLysSerPheArgHisProGlyTyrSerThrGlnThrHisValAsnAspLeuMet 60

Qy   358 CTCGTGAAGCTCAATAGCCAGGCGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGTGC 417
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    61 LeuValLeuLysLeuAsnSerGlnAlaArgLeuSerSerMetValLysLysValValArgLeuPro 80

Qy   418 TCCCGCTGCGAACCCCCTGGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGC 477
             |||||||||||||||||||||||||||||||||||||||||||||||
Db    81 SerArgCysGluProProGlyThrThrCysThrValSerGlyTrpGlyThrThrSer 100

Qy   478 CCAGATGTGACCTTTCCCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAG 537
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 ProAspValThrPheProSerAspLeuMetCysValAspValAspAsnSerProGln 120

Qy   538 GACTGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCC 597
             ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 AspCysThrLysValTyrLysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIlePro 140

Qy   598 GACTCCAAGAAAAACGCCTGCAATGGTGACTCAGGGGACCG 639
             |||||||||||||||||||||||||||||||||||||||||
Db   141 AspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyPro 154
```

```
RESULT 15
US-09-261-416-7
; Sequence 7, Application US/09261416A
; Patent No. 6291663
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Underwood, Lowell J.
; TITLE OF INVENTION: TADG-12: A No. 6291663e1 Transmembrane Serine Protease
; TITLE OF INVENTION: Overexpressed in Ovarian Carcinoma
; FILE REFERENCE: D6192
; CURRENT APPLICATION NUMBER: US/09/261,416A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 14
; SEQ ID NO 7
;   LENGTH: 154
;   TYPE: PRT
;   ORGANISM: Unknown
;   FEATURE:
;   OTHER INFORMATION: Serine protease catalytic domain of stratum corneum
;   OTHER INFORMATION: chymotryptic enzyme (Scce) homologous to similar
;   OTHER INFORMATION: domain in TADG-12
US-09-261-416-7

Alignment Scores:
Pred. No.:               9.55e-76        Length:         154
```

Search completed: July 14, 2005, 17:34:56
Job time : 43.5 secs

This Page Blank (uspto)

OM nucleic - protein search, using frame_plus_n2p model

Copyright (c) 1993 - 2005 Compugen Ltd.

GenCore version 5.1.6

Run on: July 14, 2005, 17:32:06 ; Search time 169 Seconds
(without alignments)
4430.247 Million cell updates/sec

Title: US-09-905-083A-30
Perfect score: 1780
Sequence: 1 ggattccgggctccatggc........aagaaacacaaaccctcag 969

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop   6.0 , Delext   7.0

Searched: 1726220 seqs, 386332138 residues

Total number of hits satisfying chosen parameters: 3452440

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO spool_p/US09905083/runat_14072005_140356_10721/app query.fasta_1.1159
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPEXT=0 -LOOPCLI=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09905083 @CGN_1_1_199_@runat_14072005_140356_10721
-NCPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*
1: /cgn2_6/ptcodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptcodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptcodata/2/pubpaa/US09C_PUBCOMB.pep.*
4: /cgn2_6/ptcodata/2/pubpaa/US06_NEW_PUB.pep.*
5: /cgn2_6/ptcodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptcodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptcodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptcodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptcodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptcodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptcodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptcodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptcodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptcodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptcodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptcodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptcodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptcodata/2/pubpaa/US08_PUBCOMB.pep.*
19: /cgn2_6/ptcodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptcodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptcodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptcodata/2/pubpaa/US60_PUBCOMB.pep.*

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1364	76.6	253	9	US-09-888-615-98	Sequence 98, Appl
2	1364	76.6	253	9	US-09-764-762-3	Sequence 3, Appli
3	1364	76.6	253	14	US-10-071-214-2	Sequence 2, Appli
4	1364	76.6	253	14	US-10-071-214-48	Sequence 48, Appl
5	1364	76.6	253	14	US-10-264-288-90	Sequence 90, Appl
6	1364	76.6	253	15	US-10-295-027-498	Sequence 498, App
7	1364	76.6	253	15	US-10-173-999-48	Sequence 48, Appl
8	1364	76.6	253	16	US-10-408-765A-639	Sequence 639, App
9	1364	76.6	253	16	US-10-643-795A-95	Sequence 95, Appl
10	1364	76.6	253	17	US-10-948-518-95	Sequence 95, Appl
11	1364	76.6	253	17	US-10-868-490A-1	Sequence 1, Appli
12	1350	75.8	250	15	US-10-262-511-92	Sequence 92, Appl
13	1337	75.1	257	15	US-10-344-394-38	Sequence 38, Appl
14	1252	70.3	247	15	US-10-262-511-102	Sequence 102, App
15	1252	70.3	252	15	US-10-262-511-94	Sequence 94, Appl
16	1235	69.4	225	15	US-10-600-187-4	Sequence 4, Appli
17	1230	69.1	224	11	US-09-789-210-33	Sequence 33, Appl
18	1119	62.9	224	15	US-10-262-511-104	Sequence 104, App
19	1085.5	61.0	249	14	US-10-071-214-47	Sequence 47, Appl
20	1035.5	58.2	198	15	US-10-262-511-96	Sequence 96, Appl
21	1011.5	56.8	249	14	US-10-071-214-50	Sequence 50, Appl
22	996.5	56.0	243	14	US-10-071-214-46	Sequence 46, Appl
23	940.5	52.8	226	14	US-10-071-214-49	Sequence 49, Appl
24	937.5	52.7	181	15	US-10-262-511-98	Sequence 98, Appl
25	773	43.4	144	9	US-09-796-294-4	Sequence 4, Appli
26	773	43.4	144	14	US-10-461-787-4	Sequence 4, Appli
27	614	34.5	260	9	US-09-796-294-8	Sequence 8, Appli
28	614	34.5	260	14	US-10-461-787-8	Sequence 8, Appli
29	614	34.5	260	15	US-10-072-012-609	Sequence 609, App
30	612.5	34.4	293	9	US-09-739-907-82	Sequence 82, Appl
31	612.5	34.4	293	9	US-09-739-907-97	Sequence 97, Appl
32	612.5	34.4	293	9	US-09-989-722-309	Sequence 309, App
33	612.5	34.4	293	9	US-09-989-723-309	Sequence 309, App
34	612.5	34.4	293	9	US-09-989-279-309	Sequence 309, App
35	612.5	34.4	293	9	US-09-989-727-309	Sequence 309, App
36	612.5	34.4	293	9	US-09-989-731-309	Sequence 309, App
37	612.5	34.4	293	9	US-09-989-732-309	Sequence 309, App
38	612.5	34.4	293	9	US-09-991-073-309	Sequence 309, App
39	612.5	34.4	293	9	US-09-990-442-309	Sequence 309, App
40	612.5	34.4	293	9	US-09-991-163-309	Sequence 309, App
41	612.5	34.4	293	9	US-09-993-604-309	Sequence 309, App
42	612.5	34.4	293	9	US-09-990-456-309	Sequence 309, App
43	612.5	34.4	293	9	US-09-989-721-309	Sequence 309, App
44	612.5	34.4	293	9	US-09-992-598-309	Sequence 309, App
45	612.5	34.4	293	9	US-09-989-293A-309	Sequence 309, App

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
US-09-888-615-98
; Sequence 98, Application US/0988615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 98

; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-98

Alignment Scores:
Pred. No.:
Score: 1.16e-118
Percent Similarity: 1364.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 76.63%

Length: 253
Matches: 253
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 9

US-09-905-083A-30 (1-969) x US-09-888-615-98 (1-253)

```
Qy     16 ATGGCAAGATCCCTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy     76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCC 135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AlaGlyGluGluAlaGlnValThrArgLeuLeuMetAlaProCysAlaArgGlySer 40

Qy    136 CACCCATGCAGGTGGCCCTGCTCAGCTGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 HisProTrpGlnValAlaLeuLeuSerCysAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy    196 AATGAGGCCGGCTGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AsnGluAlaTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80

Qy    256 GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTGAAGTCATTCCGC 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysPheArg 100

Qy    316 CACCCCGGCTACTCCACAGACACCATGTTAATGACCTCGTGAAGCTCAATAGC 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy    376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCGGCTGGCGAACCCCCT 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy    436 GGAACCACCTGTACTGTCTCCGGGCGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160

Qy    496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCAGGACTGACGAAGGTTAC 555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180

Qy    556 AAGGACTTACTGGAAAATTCATGCTGTGGCTGGCATCCCCGACTCCAAGAAAAACGCC 615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200

Qy    616 TGCAATGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTTGGTGTCC 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy    676 TGGGGAACTTTCCCTTGGGCGTGGCCCCTGCTCAGTGCAATGACCCAGAGTCTACACTCAAGTGTGCAAG 735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 TrpGlyThrPhePheProGlyTyrGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy    736 TTCACCAAGTGATAAATGACACCCATGAAAAAGCATCGC 774
          ||||||||||||||||||||||||||||||||||||
Db    241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

RESULT 2
US-09-764-762-3
; Sequence 3, Application US/09764762
; Patent No. US20020068341A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.

; APPLICANT: Lal, Preeti
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/764,762
; FILING DATE: 16-Jan-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/210,084
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0252 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 532504
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-764-762-3

Alignment Scores:
Pred. No.:
Score: 1.16e-118
Percent Similarity: 1364.00
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 76.63%

Length: 253
Matches: 253
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 9

US-09-905-083A-30 (1-969) x US-09-764-762-3 (1-253)

```
Qy     16 ATGGCAAGATCCCTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy     76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCC 135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AlaGlyGluGluAlaGlnValThrArgLeuLeuMetAlaProCysAlaArgGlySer 40

Qy    136 CACCCATGCAGGTGGCCCTGCTCAGCTGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy    196 AATGAGGCCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AsnGluAlaTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80

Qy    256 GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTGAAGTCATTCCGC 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysPheArg 100

Qy    316 CACCCCGGCTACTCCACAGACACCATGTTAATGACCTCGTGAAGCTCAATAGC 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy    376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGGCTGCGAACCCCCT 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy    436 GGAACCACCTGTACTGTCTCCGGGCGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160

Qy    496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCAGGACTGACGAAGGTTAC 555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180

Qy    556 AAGGACTTACTGGAAAATTCATGCTGTGGCTGGCATCCCCGACTCCAAGAAAAACGCC 615
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleLeuProAspSerLysLysAsnAla 200

Qy    616 GCGAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGTCTGGTGTCC 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 AlaGlyGluGluAlaGlnValThrArgLeuLeuMetAlaProCysAlaArgGlySer 220

Qy    736 GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCATTCCGC 315
          |||||||||||||||||||||||||||||||||||||||||||||
Db    221 TrpGlyThrPhePheProGlyTyrGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy    256 GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTCAGAAGTCCTCGAAGTCATTCCGC 315
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysPheArg 100

Qy    316 CACCCCGGCTACTCCACAGACACCATGTTAATGACCTCGTGAAGCTCAATAGC 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
```

Db   41 HisProTrpGlnValAlaAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValGlyVal 60
Qy 196 AATGAGCGGTGGTGCTCACTGGCCCCCACTGCAGAGATGAATGAGAGTGACTACACCGTGCACCTG 255
Db   61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluThrValHisLeu 80
Qy 256 GGCAGTGATACGCTGGGCGACAGGAGAGTCAGAGGATCAAGGCCTGAAGTCATTCCGC 315
Db   81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysPheArg 100
Qy 316 CACCCCGGCTACTCCACCAGACCATGTTAATGACCTCTGCGAGCTCAATAGC 375
Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
Qy 376 CAGGCCAGGCTGCATCATGGTGAAGAAGTCAGGCTGCCTCCCGCTGCGAACCCCCT 435
Db 121 GlnAlaArgLeuSerMetValLysLysValArgLeuProSerArgCysGluProPro 140
Qy 436 GGAACCACCTGTACTGTCTCCGGCTGGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC 495
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160
Qy 496 TCTGACCTCATGTGCGTGATGTCAAGCTCACGCTCATCTCCCCCAGGACTGCGACGAGGTTTAC 555
Db 161 SerAspLeuMetCysValMetSerSerProGlnAspCysThrLysValTyr 180
Qy 556 AAGGACTTACTGGAAATTCCATGCTGTGCGCTGGCATCCCGACTCCCAAGAAAAACGCC 615
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
Qy 616 TGCAATGGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGTGTCC 675
Db 201 CysAsnGlyAspSerGlyAspProValCysArgGlyThrLeuGlnGlyValSer 220
Qy 676 TGGGGAACTTCCCTGCCGCCAACCCAATGACCCAGGAGTCTACACTCCAAGTGTGCAAG 735
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240
Qy 736 TTCACCAAGTGGATAAATGACACCATGAAAAGCATCGC 774
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 4
US-10-071-214-48
; Sequence 48, Application US/10071214
; Publication No. US2003006609A1
; GENERAL INFORMATION:
; APPLICANT: HANSSON, Lennart
; APPLICANT: EGELRUD, Torbjorn
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN
; FILE REFERENCE: HANSSON=3A
; CURRENT APPLICATION NUMBER: US/10/071,214
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: US 60/267,422
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: DK PA 2001 00218
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Deduced amino acid sequence from the C-terminal part of SCCE from
; OTHER INFORMATION: homo sapiens.
US-10-071-214-48

Alignment Scores:
Pred. No.:   1.16e-118   Length: 253
Score: 1364.00   Matches: 253
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0

Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120
Qy 376 CAGGCCAGGCTGCATCATGGTGAAGAAGTCAGGCTGCCTCCCGCTGCGAACCCCCT 435
Db 121 GlnAlaArgLeuSerMetValLysLysValArgLeuProSerArgCysGluProPro 140
Qy 436 GGAACCACCTGTACTGTCTCCGGCTGGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC 495
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160
Qy 496 TCTGACCTCATGTGCGTGATGTCAAGCTCACGCTCATCTCCCCCAGGACTGCGACGAGGTTTAC 555
Db 161 SerAspLeuMetCysValMetSerSerProGlnAspCysThrLysValTyr 180
Qy 556 AAGGACTTACTGGAAATTCCATGCTGTGCGCTGGCATCCCGACTCCCAAGAAAAACGCC 615
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200
Qy 616 TGCAATGGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGTGTCC 675
Db 201 CysAsnGlyAspSerGlyAspProValCysArgGlyThrLeuGlnGlyValSer 220
Qy 676 TGGGGAACTTCCCTGCCGCCAACCCAATGACCCAGGAGTCTACACTCCAAGTGTGCAAG 735
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240
Qy 736 TTCACCAAGTGGATAAATGACACCATGAAAAGCATCGC 774
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 3
US-10-071-214-2
; Sequence 2, Application US/10071214
; Publication No. US20030066099A1
; GENERAL INFORMATION:
; APPLICANT: HANSSON, Lennart
; APPLICANT: EGELRUD, Torbjorn
; TITLE OF INVENTION: SCCE MODIFIED TRANSGENIC MAMMALS AND THEIR USE AS MODELS OF HUMAN
; FILE REFERENCE: HANSSON=3A
; CURRENT APPLICATION NUMBER: US/10/071,214
; CURRENT FILING DATE: 2002-02-11
; PRIOR APPLICATION NUMBER: US 60/267,422
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: DK PA 2001 00218
; PRIOR FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-071-214-2

Alignment Scores:
Pred. No.:   1.16e-118   Length: 253
Score: 1364.00   Matches: 253
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 76.63%   Indels: 0
DB: Gaps: 14

US-09-905-083A-30 (1-969) x US-10-071-214-2 (1-253)

Qy   16 ATGGCCAAGATCCCTTCTCCGCCCTGCCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
Db   1 MetAlaArgSerLeuLeuProLeuGlnIleLeuLeuSerLeuAlaLeuGluThr 20
Qy   76 GCAGGAGGAGAAGCCCAGGTGACAAGAGATTATTGATGCGGCCCCATGTGCGAAGAGGCTCC 135
Db   21 AlaGlyGlyGluAlaGlnAlaGlnValAspLysLysIleIleAspAlaProCysSerAlaArgGlySer 40
Qy 136 CACCCATGGCAGTGGCCCTGCTCAGTCGCAATCAGCTCCACTGCGGAGCGCGTCCTGTC 195

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-264-283-90

Alignment Scores:
Pred. No.:   1.16e-118   Length: 253
Score: 1364.00   Matches: 253
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 76.63%   Indels: 0
DB: 14   Gaps: 0

US-09-905-083A-30 (1-969) x US-10-264-283-90 (1-253)

```
Qy   16 ATGGCAAGATCCCTTCCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT  75
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr  20

Qy   76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCC  135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21  AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer  40

Qy  136 CACCCATGGCAGGTGGCCCTGCTCCAGTGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC  195
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41  HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal  60

Qy  196 AATGAGCGCGTGGGTGCTCACTGCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG  255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu  80

Qy  256 GGCAGTGATACGCTGGGCACAGGAGAGCTCAGGAGGCCTCAGAAGTCATTCCGC  315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81  GlySerAspThrLeuGlyThrGlyGluLeuArgArgAlaGlnArgIleLysSerPheArg  100

Qy  316 CACCCCGGCTACTCCACACAGACCCCATGTTAATGACCTCGTGAAGCTCAATAGC  375
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101  HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer  120

Qy  376 CAGGCCAGGCTGTCATCCATGGTGAAGAAGTCAGGCTGCCTCCGCTGCGAACCCCCT  435
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro  140

Qy  436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCAGCCCAGATGTGACCTTTCCC  495
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141  GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro  160

Qy  496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCGAAGGTTTAC  555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161  SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr  180

Qy  556 AAGGACTTACTGGAAATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAACGCC  615
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla  200

Qy  616 TGCAATGTGACTCAGGGGGACCGTTGGTGTGCAGAGGGTACCCTGCAGAGGTTTGTGTCC  675
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201  CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer  220

Qy  676 TGGGGAACTTCCCTTGCGGCCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG  735
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221  TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys  240

Qy  736 TTCACCAAGTGGATAAATGACACCATGAAAAGCATCGC  774
          ||||||||||||||||||||||||||||||||||||
Db  241  PheThrLysTrpIleAsnAspThrMetLysLysHisArg  253
```

Query Match: 76.63%   Indels: 0
DB: 14   Gaps: 0

US-09-905-083A-30 (1-969) x US-10-071-214-48 (1-253)

```
Qy   16 ATGGCAAGATCCCTTCCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT  75
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1  MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr  20

Qy   76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCC  135
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21  AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer  40

Qy  136 CACCCATGGCAGGTGGCCCTGCTCCAGTGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC  195
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41  HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal  60

Qy  196 AATGAGCGCGTGGGTGCTCACTGCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG  255
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61  AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu  80

Qy  256 GGCAGTGATACGCTGGGCACAGGAGAGCTCAGGAGGCCTCGAAGTCATTCCGC  315
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81  GlySerAspThrLeuGlyThrGlyGluLeuArgArgAlaGlnArgIleLysSerPheArg  100

Qy  316 CACCCCGGCTACTCCACACAGACCCCATGTTAATGACCTCGTGAAGCTCAATAGC  375
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101  HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer  120

Qy  376 CAGGCCAGGCTGTCATCCATGGTGAAGAAGTCAGGCTGCCTCCGCTGCGAACCCCCT  435
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121  GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro  140

Qy  436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCAGCCCAGATGTGACCTTTCCC  495
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141  GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro  160

Qy  496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCGAAGGTTTAC  555
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161  SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr  180

Qy  556 AAGGACTTACTGGAAATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAACGCC  615
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181  LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla  200

Qy  616 TGCAATGTGACTCAGGGGGACCGTTGGTGTGCAGAGGGTACCCTGCAGAGGTTTGTGTCC  675
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201  CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer  220

Qy  676 TGGGGAACTTCCCTTGCGGCCAACCCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG  735
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221  TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys  240

Qy  736 TTCACCAAGTGGATAAATGACACCATGAAAAGCATCGC  774
          ||||||||||||||||||||||||||||||||||||
Db  241  PheThrLysTrpIleAsnAspThrMetLysLysHisArg  253
```

; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 498
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-498

Alignment Scores:
Pred. No.:
Score: 1364.00   Length: 253
Percent Similarity: 100.00%   Matches: 253
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 76.63%   Mismatches: 0
DB: 15   Indels: 0
                                              Gaps: 0

US-09-905-083A-30 (1-969) x US-10-295-027-498 (1-253)

Qy   16 ATGGCAAGATCCCTTCTCCTGCCCCTGCGAGATCTTACTGCTATCCTTAGCCTTGGAAACT   75
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr   20

Qy   76 GCAGGAGGAAGAAGCCCAGGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGAGGCTCC 135
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AlaGlyGluGluAlaGlnGlyAspLysIleIleIleAspGlyAlaProCysAlaArgGlySer   40

Qy 136 CACCCATGGCAGGTGGCCCCTGCTCAGTGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 HisProTrpGlnValAlaAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal   60

Qy 196 AATGAGCCGTGGGTGCTCACTGCCGCCACCTGCAAGATGAATGAGTACCGTGCACCTG 255
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu   80

Qy 256 GGCAGTGATACCGCTGGGCCACCAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC 315
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 GlySerAspThrLeuGlyLeuLeuGlyGlyAspLysArgArgArgSerGlnArgIleLysAlaSerLysLysSerPheArg 100

Qy 316 CACCCCGGCTACTCCACAGACCCATGTTAATGACCTCATGCTGTGAAGCTCAATAGC 375

; RESULT 7
; US-10-173-999-48
; Sequence 48, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-0024200US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-48

Alignment Scores:
Pred. No.:
Score: 1364.00   Length: 253
Percent Similarity: 100.00%   Matches: 253
Best Local Similarity: 100.00%   Conservative: 0
Query Match: 76.63%   Mismatches: 0
DB: 15   Indels: 0
                                              Gaps: 0

US-09-905-083A-30 (1-969) x US-10-173-999-48 (1-253)

Qy   16 ATGGCAAGATCCCTTCTCCTGCCCCTGCGAGATCTTACTGCTATCCTTAGCCTTGGAAACT   75
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy 376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCGCTGCGAACCCCT 435
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GlnAlaArgLeuSerSerMetValLysValArgLeuProSerArgCysGluProPro 140

Qy 436 GGAACCACCTGTACTGTCTCCGGCTGGGGGCACTACCAGAGCACCCAGATGTGACCTTCC 495
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro 160

Qy 496 TCTGACCTCATGTGCTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTAC 555
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180

Qy 556 AAGGACTTACTGGAAAATTCCATGCTGTGCGCTGCCATCCCGACTCCAAGAAAACGCC 615
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200

Qy 616 TGCAATGGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCC 675
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 201 CysAsnGlyAspSerGlyGlyProLeuValCysAsnGlyThrLeuGlnGlyLeuValSer 220

Qy 676 TGGGGAACTTTCCCTTGCGGCCAACCAATGACCCAGGAGTCTACACTCCAAGTGTGCAAG 735
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy 736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
             |||||||||||||||||||||||||||||||||||||||
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

US-10-408-765A-639

Alignment Scores:

Pred. No.:	
Score:	1364.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	76.63%
DB:	

Length: 253
Matches: 253
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-905-083A-30 (1-969) × US-10-408-765A-639 (1-253)

```
Qy     16 ATGGCAAGATCCCTTCCTCCGCCCTGCGAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy     76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCC 135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AlaGlyGluGluAlaGlnValAlaArgLeuLeuMetGlyProCysAlaArgGlySer 40

Qy    136 CACCCATGCGAGGTGGCCCTGCTCACTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy    196 AATGAGGCGCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80

Qy    256 GGCAGTGATACGCTGGGCGACAGAGAGTCAGGGATCAAGGCCTGAAGTCATTCCGC 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg 100

Qy    316 CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCGTGAAGCTCAATAGC 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy    376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCCT 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy    436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160

Qy    496 TCTGACCTCATTGCGTGCGGATGTCAAGCTGCTGCCCCCAGGACTGCCACGAAGGTTTAC 555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 SerAspLeuMetCysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyr 180

Qy    556 AAGGACTTACTGGAAAATTCCATGCTGTCGCGCTGGCATCCCCGACTCCCAAGAAAAACGCC 615
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProSerLysLysAsnAla 200

Qy    616 TGCAATGGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCGTGCAAGGTGTCGTGTCC 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy    676 TGGGGAACTTTCCCTTGCGGCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy    736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
          |||||||||||||||||||||||||||||||||||||||
Db    241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

RESULT 9
US-10-643-795A-95
; Sequence 95, Application US/10643795A
; Publication No. US20040241703A1
; GENERAL INFORMATION:
; APPLICANT: FREDERIC J. DESAUVAGE
; APPLICANT: GRETCHEN FRANTZ
; APPLICANT: KENNETH J. HILLAN
; APPLICANT: PAUL POLAKIS

```
Db      1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy     76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCC 135
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AlaGlyGluGluAlaGlnValAlaArgLeuLeuMetGlyProCysAlaArgGlySer 40

Qy    136 CACCCATGCGAGGTGGCCCTGCTCACTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy    196 AATGAGGCGCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG 255
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80

Qy    256 GGCAGTGATACGCTGGGCGACAGAGAGTCAGGGATCAAGGCCTGAAGTCATTCCGC 315
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg 100

Qy    316 CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCGTGAAGCTCAATAGC 375
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer 120

Qy    376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCT 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GlnAlaArgLeuSerSerMetValLysLysValArgLeuProSerArgCysGluProPro 140

Qy    436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro 160

Qy    496 TCTGACCTCATTGCGTGCGGATGTCAAGCTGCTGCCCCCAGGACTGCCACGAAGGTTTAC 555
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 SerAspLeuMetCysValAspValLysLeuLeuSerProGlnAspCysThrLysValTyr 180

Qy    556 AAGGACTTACTGGAAAATTCCATGCTGTCGCGCTGGCATCCCCGACTCCCAAGAAAAACGCC 615
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProSerLysLysAsnAla 200

Qy    616 TGCAATGGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCGTGCAAGGTGTCGTGTCC 675
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy    676 TGGGGAACTTTCCCTTGCGGCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG 735
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys 240

Qy    736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
          |||||||||||||||||||||||||||||||||||||||
Db    241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253
```

RESULT 8
US-10-408-765A-639
; Sequence 639, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 639
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Homo sapiens

Db 161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr 180

Qy 556 AAGGACTTACTGAAATTCCATGCTGTGCGCTGGCATCCCGACTCCCAAGAAAAACGCC 615
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 181 LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla 200

Qy 616 TGCAATGGTGACTCAGGGGGACCGGTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCC 675
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer 220

Qy 676 TGGGGAACTTTCCCTGCGGCCAACCCAGTGACCCAGGAGTCTACACCAGTGTGCAAG 735
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 221 TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValThrGlnValCysLys 240

Qy 736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC 774
         |||||||||||||||||||||||||||||||||||||
Db 241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg 253

RESULT 10
US-10-948-518-95
; Sequence 95, Application US/10948518
; Publication No. US20050064492A1
; GENERAL INFORMATION:
;   APPLICANT: FREDERIC J. DESAUVAGE
;   APPLICANT: GRETCHEN FRANTZ
;   APPLICANT: KENNETH J. HILLAN
;   APPLICANT: PAUL POLAKIS
;   APPLICANT: ANDREW POLSON
;   APPLICANT: VICTORIA SMITH
;   APPLICANT: SUSAN D. SPENCER
;   APPLICANT: THOMAS D. WU
;   APPLICANT: ZEMIN ZHANG
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;   TITLE OF INVENTION: TREATMENT OF TUMOR
;   FILE REFERENCE: P5026R1-US
;   CURRENT APPLICATION NUMBER: US/10/948,518
;   CURRENT FILING DATE: 2004-09-22
;   PRIOR APPLICATION NUMBER: US/10/643,795
;   PRIOR FILING DATE: 2003-08-19
;   PRIOR APPLICATION NUMBER: US 60/404,809
;   PRIOR FILING DATE: 2002-08-19
;   PRIOR APPLICATION NUMBER: US 60/405,645
;   PRIOR FILING DATE: 2002-08-21
;   PRIOR APPLICATION NUMBER: US 60/413,192
;   PRIOR FILING DATE: 2002-09-23
;   PRIOR APPLICATION NUMBER: US 60/419,008
;   PRIOR FILING DATE: 2002-10-15
;   PRIOR APPLICATION NUMBER: US 60/426,847
;   PRIOR FILING DATE: 2002-11-15
;   PRIOR APPLICATION NUMBER: US 60/484,959
;   PRIOR FILING DATE: 2003-07-02
;   NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 95
;   LENGTH: 253
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-948-518-95

Alignment Scores:
Pred. No.:   Length: 253
Score: 1.16e-118   Matches: 253
Percent Similarity: 1364.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 76.63%   Gaps: 0
                    17

US-09-905-083A-30 (1-969) x US-10-948-518-95 (1-253)

Qy   16 ATGGCAAGATCCCTTCTCCTGCCCCTGCCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy   76 GCAGGAGAGAAGCCCAGGTGACAAGATTATTGATGGGCGCCCCATGTGCAAGAGGCTCC 135
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AlaGlyGluGluLysProLeuGlnIleIleAspGlyArgProMetCysLysArgLeuSer 40

RESULT 10
US-10-643-795A-95
; Sequence 95, Application US/10643795A
; Publication No. US20040138420A1
; GENERAL INFORMATION:
;   APPLICANT: ANDREW POLSON
;   APPLICANT: VICTORIA SMITH
;   APPLICANT: SUSAN D. SPENCER
;   APPLICANT: THOMAS D. WU
;   APPLICANT: ZEMIN ZHANG
;   TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DIAGNOSIS AND
;   TITLE OF INVENTION: TREATMENT OF TUMOR
;   FILE REFERENCE: P5026R1-US
;   CURRENT APPLICATION NUMBER: US/10/643,795A
;   CURRENT FILING DATE: 2003-08-19
;   PRIOR APPLICATION NUMBER: US 60/404,809
;   PRIOR FILING DATE: 2002-08-19
;   PRIOR APPLICATION NUMBER: US 60/405,645
;   PRIOR FILING DATE: 2002-08-21
;   PRIOR APPLICATION NUMBER: US 60/413,192
;   PRIOR FILING DATE: 2002-09-23
;   PRIOR APPLICATION NUMBER: US 60/419,008
;   PRIOR FILING DATE: 2002-10-15
;   PRIOR APPLICATION NUMBER: US 60/426,847
;   PRIOR FILING DATE: 2002-11-15
;   PRIOR APPLICATION NUMBER: US 60/484,959
;   PRIOR FILING DATE: 2003-07-02
;   NUMBER OF SEQ ID NOS: 158
; SEQ ID NO 95
;   LENGTH: 253
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-10-643-795A-95

Alignment Scores:
Pred. No.:   Length: 253
Score: 1.16e-118   Matches: 253
Percent Similarity: 1364.00   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 100.00%   Indels: 0
DB: 76.63%   Gaps: 0
                    16

US-09-905-083A-30 (1-969) x US-10-643-795A-95 (1-253)

Qy   16 ATGGCAAGATCCCTTCTCCTGCCCCTGCCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy   76 GCAGGAGAGAAGCCCAGGTGACAAGATTATTGATGGGCGCCCCATGTGCAAGAGGCTCC 135
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 AlaGlyGluGluGlnLysProLeuGlnIleLeuIleAspLysIleAspAspGlyAlaArgGlySer 40

Qy 136 CACCCATGGCAGGTGGCCCTGCTCAGTGCCAATCAGCTCCACTGCGGAGGCGTCCTGGTC 195
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal 60

Qy 196 AATGAGCGGCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG 255
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 AsnGluArgTrpValLeuThrAlaLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu 80

Qy 256 GGCAGTGATACGCTGGGCGACCAGGAGACTCAGAGGATCAAGGCCTCGAAGTCATTCCGC 315
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 GlySerAspThrLeuGlyAspGlnArgThrGlnArgIleLysAlaSerLysSerPheArg 100

Qy 316 CACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCGTGAAGCTCAATAGC 375
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuValLysLeuAsnSer 120

Qy 376 CAGGCCAGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTGCCCGTTGCGAACCCCCT 435
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 121 GlnAlaArgLeuSerSerMetCysHisProTrpValLysLysValArgLeuProCysProPro 140

Qy 436 GGAACCACCTGTACTGTCTCCGGTGGGCACTCACCACCAGACGATGTGACCCTTTCC 495
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 141 GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrProPro 160

Qy 496 TCTGACCTCATGTGTGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTAC 555
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

DB: 17   Gaps: 0

US-09-905-083A-30 (1-969) x US-10-868-490A-1 (1-253)

```
Qy     16  ATGGCAAGATCCCTTCCTCGCCCCTGCCCAGATCTTACTGCTGCTATCCTTAGCCTTGAAACT  75
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MetAlaArgSerLeuLeuLeuProLeuProGlnIleLeuLeuLeuLeuSerLeuAlaLeuGluThr  20

Qy     76  GCAGGAGAAGCCCAGGTGACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCC  135
           ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21  AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer  40

Qy    136  CACCCATGGCAGGTGGCCCCTGCTCAGTGGCAATCAGCTCCACTGCGAGGCGTCCTGGTC  195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41  HisProTrpGlnValAlaLeuLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal  60

Qy    196  AATGAGCCGCTGGGTGCTCACTGCCGCCCACTGAACAAGATGAATAGTACACCGTGCACCTG  255
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu  80

Qy    256  GGCAGTGATACGCTGGGCACAGAGAGCTCAGAGACTCGAAGCTCAATAGC  315
           ||||||||||||||||||||||||||||||||||||||||||||||
Db     81  GlySerAspThrLeuGlyThrGluSerSerGluThrArgSerSerPheArg  100

Qy    316  CACCCCGGCTACTCCACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC  375
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    101  HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer  120

Qy    376  CAGGCCAGGCTGTCATCCATGGTCAGGGCCACTACCACAGATGTGACCTTTCCC  435
           ||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GlnAlaArgLeuSerSerMetValLysValArgAlaArgIleLeuLysAlaSerPheArg  140

Qy    436  GGAACCACCTGTACTGTCTCCGGCTGGGCACTCCACCAGAGTGGGGCACTCCACCAGA  495
           |||||||||||||||||||||||||||||||||||||||||||||||
Db    141  GlyThrThrCysThrValSerGlyTrpGlyThrThrSerProAspValThrPhePro  160

Qy    496  TCTGACCTTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTCGACGAAGGTTTAC  555
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    161  SerAspLeuMetCysValAspValLysLeuIleSerProArgLeuAspProGlnAspCysThrLysValTyr  180

Qy    556  AAGGACTTACTGGAAAATTCCATGGTCGCGTGGCATCCCCGACTCCCAAGAAAAACGCC  615
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    181  LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla  200

Qy    616  TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGTGTCC  675
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db    201  CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer  220

Qy    676  TGGGGAACTTCCCTGCGGCCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG  735
           ||||||||||||||||||||||||||||||||||||||||||||||
Db    221  TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys  240

Qy    736  TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC  774
           |||||||||||||||||||||||||||||||||||
Db    241  PheThrLysTrpIleAsnAspThrMetLysLysHisArg  253
```

(Second alignment — left portion)

```
Db     21  AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer  40

Qy    136  CACCCATGGCAGGTGGCCCCTGCTCAGTGGCAATCAGCTCCACTGCGAGGCGTCCTGGTC  195
Db     41  HisProTrpGlnValAlaLeuLeuLeuSerGlyAsnGlnLeuHisCysGlyValLeuVal  60

Qy    196  AATGAGCCGCTGGGTGCTCACTGCCGCCCACTGAACAAGATGAATAGTACACCGTGCACCTG  255
Db     61  AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu  80

Qy    256  GGCAGTGATACGCTGGGCACAGAGAGCTCAGAGACTCGAAGCTCAATAGC  315
Db     81  GlySerAspThrLeuGlyThrGluSerSerGluThrArgSerSerPheArg  100

Qy    316  CACCCCGGCTACTCCACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGC  375
Db    101  HisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSer  120

Qy    376  CAGGCCAGGCTGTCATCCATGGTCAGGGCCACTACCACAGATGTGACCTTTCCC  435
Db    121  GlnAlaArgLeuSerSerMetValLysValArgAlaArgIleLeuLysAlaSerPheArg  140

Qy    436  GGAACCACCTGTACTGTCTCCGGCTGGGCACTCCACCAGA  495
Db    141  GlyThrThrCysThrValSerGlyTrpGlyThr  160

Qy    496  TCTGACCTTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTCGACGAAGGTTTAC  555
Db    161  SerAspLeuMetCysValAspValLysLeuIleSerPro  180

Qy    556  AAGGACTTACTGGAAAATTCCATGGTCGCGTGGCATCCCCGACTCCCAAGAAAAACGCC  615
Db    181  LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla  200

Qy    616  TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGTGTCC  675
Db    201  CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer  220

Qy    676  TGGGGAACTTCCCTGCGGCCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAG  735
Db    221  TrpGlyThrPheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys  240

Qy    736  TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC  774
Db    241  PheThrLysTrpIleAsnAspThrMetLysLysHisArg  253
```

Alignment Scores:
Pred. No.:   1
Score: 1364.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 76.63%

Length: 253
Matches: 253
Conservative: 0
Mismatches: 0
Indels: 0

(Expect value) 1.16e-118

```
;   APPLICANT:   Ort, Tatiana
;   APPLICANT:   Gorman, Linda
;   APPLICANT:   Zerhusen, Bryan D.
;   APPLICANT:   Anderson, David W.
;   APPLICANT:   Zhong, Mei
;   APPLICANT:   Catterton, Elina
;   APPLICANT:   Ji, Weizhen
;   APPLICANT:   Miller, Charles E.
;   APPLICANT:   Rastelli, Luca
;   APPLICANT:   Stone, David J.
;   APPLICANT:   Pena, Carol E. A.
;   APPLICANT:   Shenoy, Suresh G.
;   APPLICANT:   Shimkets, Richard A.
;   APPLICANT:   Rothenberg, Mark E.
;   APPLICANT:   Leach, Martin D.
;   APPLICANT:   Agee, Michele L.
;   APPLICANT:   Berghs, Constance.
;   TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
;   FILE REFERENCE: 21402-462C
;   CURRENT APPLICATION NUMBER: US/10/262,511
;   CURRENT FILING DATE: 2003-05-28
;   PRIOR APPLICATION NUMBER: 60/326,483
;   PRIOR FILING DATE: 2001-10-02
;   PRIOR APPLICATION NUMBER: 60/373,815
;   PRIOR FILING DATE: 2002-04-19
;   PRIOR APPLICATION NUMBER: 60/327,917
;   PRIOR FILING DATE: 2001-10-09
;   PRIOR APPLICATION NUMBER: 60/381,642
;   PRIOR FILING DATE: 2002-05-17
;   PRIOR APPLICATION NUMBER: 60/328,029
;   PRIOR FILING DATE: 2002-10-09
;   PRIOR APPLICATION NUMBER: 60/381,038
;   PRIOR FILING DATE: 2002-05-16
;   PRIOR APPLICATION NUMBER: 60/328,056
;   PRIOR FILING DATE: 2001-10-09
;   PRIOR APPLICATION NUMBER: 60/373,260
;   PRIOR FILING DATE: 2002-04-17
;   PRIOR APPLICATION NUMBER: 60/373,826
;   PRIOR FILING DATE: 2002-04-19
;   PRIOR APPLICATION NUMBER: 60/327,435
;   PRIOR FILING DATE: 2001-10-05
;   Remaining Prior Application data removed - See File Wrapper or PALM.
;   NUMBER OF SEQ ID NOS: 439
;   SOFTWARE: CuraSeqList version 0.1

SEQ ID NO 92
;   LENGTH: 250
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-262-511-92

Alignment Scores:
Pred. No.:
Score:                    2.37e-117
                          1350.00
Percent Similarity:       100.00%        Length:          250
Best Local Similarity:    100.00%        Matches:         250
Query Match:              75.84%         Conservative:    0
DB:                       15             Mismatches:      0
                                         Indels:          0
                                         Gaps:            0

US-09-905-083A-30 (1-969) x US-10-262-511-92 (1-250)

Qy   25 TCCCTTCTCCTGCCCCTGCCAGATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAA 84
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1 SerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGlyThrAlaGlyGlu 20

Qy   85 GAAGCCCAGGGTGACAAGAGATTATTGATGGCGCCGCCCATGTGCAAGAGGCTCCACCCATGG 144
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   21 GluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySerHisProTrp 40

Qy  145 CAGGTGGCCCTGCTCAGTGCAATCAGCTCCACTGCCGAGGCGCTCCTGGTCAATGAGCGC 204
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   41 GlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArg 60

Qy  205 TGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGAT 264
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   61 TrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeuGlySerAsp 80

Qy  265 ACGTGGGCGACAGGAGAGCTCAGAGATCAAGGACTCGAAGTCATTCCGACCCCCGGC 324
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   81 ThrLeuGlyAspArgArgAlaGlnArgIleLysSerPheArgLysAlaSerAsp 100

Qy  325 TACTCCACAGACCCATGTTAATGACCTCGTGCAGAGCTCAATAGCCAGGCCAGG 384
            |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  101 TyrSerThrGlnThrHisValAsnAspLeuValGlnSerSerGlnAlaArg 120

Qy  385 CTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTGCCCCGCTGCGAACCCCTGGAACCACC 444
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  121 LeuSerSerMetValLysLysValArgLeuProSerProGluProLeuProGlyThrThr 140

Qy  445 TGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCCTCGACCTC 504
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  141 CysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPheProSerAspLeu 160

Qy  505 ATGTGCGTGGATGTCAAGCTCATCTCCCCAGGACTGCACGAGAGGTTTACAAGGACTTA 564
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  161 MetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyrLysAspLeu 180

Qy  565 CTGGAAAATTCCATGCTGCGCTGGCATCCCGACTCCCAAGAAAAACGCCTGCAATGGT 624
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  181 LeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAlaCysAsnGly 200

Qy  625 GACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCCTGGGAACT 684
            |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  201 AspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThr 220

Qy  685 TTCCCTTGCGGCCAACCCAATGACCACCAGGAGTCACACTCCAAGTGTGCAAGTTCACCAAG 744
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  221 PheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLysPheThrLys 240

Qy  745 TGGATAAATGACACCATGAAAAAGCATCGC 774
            ||||||||||||||||||||||||||||||
Db  241 TrpIleAsnAspThrMetLysLysHisArg 250


RESULT 13
US-10-344-394-38
;   Sequence 38, Application US/10344394
;   Publication No. US2004005342A1
GENERAL INFORMATION:
;   APPLICANT: Yousef, George M.
;   APPLICANT: Diamandis, Eleftherios P.
;   TITLE OF INVENTION: NOVEL KALLIKREIN GENE
;   FILE REFERENCE: 11757.51USWO
;   CURRENT APPLICATION NUMBER: US/10/344,394
;   CURRENT FILING DATE: 2003-02-11
;   PRIOR APPLICATION NUMBER: PCT/CA01/01141
;   PRIOR FILING DATE: 2001-08-10
;   PRIOR APPLICATION NUMBER: US 60/224,853
;   PRIOR FILING DATE: 2000-08-11
;   NUMBER OF SEQ ID NOS: 50
;   SOFTWARE: PatentIn version 3.1
SEQ ID NO 38
;   LENGTH: 257
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-344-394-38

Alignment Scores:
Pred. No.:
Score:                    3.94e-116
                          1337.00
Percent Similarity:       97.67%         Length:          257
Best Local Similarity:    97.67%         Matches:         252
Query Match:              75.11%         Conservative:    0
DB:                       15             Mismatches:      0
                                         Indels:          6
                                         Gaps:            2

US-09-905-083A-30 (1-969) x US-10-344-394-38 (1-257)

Qy   16 ATGGCAAGATCCCTTCTCCTGCCCCTGCCAGATCTTACTGCTATCCTTACTGCCTTGGAAACT 75
```

Db |||||||||||||||||||||||||||||||||||||||
1 MetAlaArgSerLeuLeuProLeuGlnIleLeuLeuLeuLeuSerLeuAlaLeuGluThr 20

Qy GCAGGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGAGGCTCC 135
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db AlaGlyGluGluAlaGlnValThrArgLeuLeuMetAlaProCysLysArgGluSer 40
21 AlaGlyGluGluAlaGlnGlyAspLysIleIleIleAspGlyAlaProCysAlaArgGlySer

Qy CACCCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGCTCCACTGC-------------- 180
   |||||||||||||||||||||||||||||||||||||||||||
Db HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysHisSerCysCysGlu 60
41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysHisSerCysCysGlu

Qy GGAGGCGTCCTGTCAATAGAGCGCTGGGTGCTCACTGCGCCGCCACTGCCAAGATGAATGAG 240
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GlyGlyValLeuValAlaAsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGlu 80
61 GlyGlyValLeuValAlaAsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGlu

Qy TACACCGTGCACCTGGGCAGTGACTACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCC 300
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db TyrThrValHisLeuGlySerAspTyrAlaGlyArgGlnGluSerSerGluAspArgArg 100
81 TyrThrValHisLeuGlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaAla

Qy TCGAAGTCATTCCGCCACCCCGGCTACTCCACACAGACCCATGTTAATGACCTCATGCTC 360
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db SerLysSerPheArgHisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeu 120
101 SerLysSerPheArgHisProGlyTyrSerThrGlnThrHisValAsnAspLeuMetLeu

Qy GTGAAGCTCAATAGCCAGGCCAGGCTGTCATCATGGTGAAGAAGTCAGGCTGCCCTCC 420
   |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ValLysLeuAsnSerGlnAlaArgLeuSerSerMetValLysLysValArgLeuProSer 140
121 ValLysLeuAsnSerGlnAlaArgLeuSerSerMetValLysLysValArgLeuProSer

Qy CGCTGCGAACCCCTGGAACCACCTGTCTCCGGCTGGGGCACTACCACGAGCCCA 480
   |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ArgCysGluProProGlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerPro 160
141 ArgCysGluProProGlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerPro

Qy GATGTGACCTTTCCCTCTGACCTCATGTGCGTGGATGTCAAGTCAATCTCCCCCAGGAC 540
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db AspValThrPhePro---AspLeuMetCysValAspValLysLeuIleSerProGlnAsp 179
161 AspValThrPhePro---AspLeuMetCysValAspValLysLeuIleSerProGlnAsp

Qy TGCACGAAGGTTTACAGGGAACTTTCCCTTGCGGCCCGGTGGCATCCCCGAC 600
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db CysThrLysValTyrLysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleLeuProAsp 199
180 CysThrLysValTyrLysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleLeuProAsp

Qy TCCAAGAAAAACGCCTGCAATGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTG 660
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db SerLysLysAsnAlaCysAsnValThrGlnGlyAspArgSerGlyGlyProLeuValCysAspGlyValLysArgGlyThrLeu 219
200 SerLysLysAsnAlaCysAsnValThrGlnGlyAspSerGlyGlyProLeuValCysAspGlyValLysArgGlyThrLeu

Qy CAAGGTCTGGTGTCCTGGGGAACTTTCCCTTGCGGCCAACCCAATGACCCAGGAGTCTAC 720
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db GlnGlyLeuValSerTrpGlyThrPheGlyCysGlyGlnProAsnAspProGlyValTyr 239
220 GlnGlyLeuValSerTrpGlyThrPheGlyCysGlyGlnProAsnAspProGlyValTyr

Qy ACTCAAGTGTGCAAGTTCACCAAGTCGGATAAATGACACCATGAAAAAGCATCGC 774
   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ThrGlnValCysLysPheThrLysTrpIleAsnAspThrMetLysLysHisArg 257
240 ThrGlnValCysLysPheThrLysTrpIleAsnAspThrMetLysLysHisArg

US-09-905-083A-30 (1-969) x US-10-262-511-102 (1-247)

Qy GCCCAGGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGAGGCTCCCACCCATGGCAG 147
88 |||||::::||||||||||||||||||||||||||||||||||||||||||||
Db AlaArgGlyAspLysIleIleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGln 30
11

Qy GTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTCAATGAGCGCTGG 207
148 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ValAlaLeuLeuSerGlyAsnGlnLeuHisCysHisSerCysCysGluGlyGlyValLeuValAlaAsnGluArgTrp 50
31

Qy GTGCTCACTGCGCCGCCACTGCCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACG 267
208 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db ValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeuGlySerAspThr 70
51

Qy CTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGCCACCCCGGCTAC 327
268 |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db LeuGlyAspArgArgAlaGlnArgIleLysAlaAlaSerLysSerPheArgHisProGlyTyr 90
71

Qy TCCACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGGCCAGGCTG 387
328 |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db   91 SerThrGlnThrHisAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeu 110

Qy   388 TCATCCATGGTGAAGAAAGTCAGGCTGCCTCCGCTGCGAACCCCTGGAACCACCTGT 447
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   111 SerSerMetValLysLysValLysArgLeuProSerArgCysGluProProGlyThrThrCys 130

Qy   448 ACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTCCCTCGACCTCATG 507
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   131 ThrValSerGlyTrpGlyThrThrThrSerProAspValThrPheProSerAspLeuMet 150

Qy   508 TGGGTGGATGTCAAGCTCATCTCCCCAGGACTGCACGAAGGTTACAGGACTTACTG 567
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   151 CysValAspValLysLeuIleSerProGlnAspCysThrLysGlyThrLysTyrLysAspLeuLeu 170

Qy   568 GAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAACGCCTGCAATGGTGAC 627
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   171 GluAsnSerMetLeuCysAlaGlyIleProThrProLysLysAsnAlaCysAsnGlyAsp 190

Qy   628 TCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCCTGGGAACTTTC 687
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   191 SerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPhe 210

Qy   688 CCTTGCGGCAACCCAATGACCCAGGAGTCTCACTCAAGTGTGCAAGTTCACCAAGTGG 747
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   211 ProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrp 230

Qy   748 ATAAATGACACCATGAAAAAGCATCGC 774
          |||||||||||||||||||||||||||
Db   231 IleAsnAspThrMetLysLysHisArg 239

RESULT 15
US-10-262-511-94
; Sequence 94, Application US/10262511
; Publication No. US20040038223A1
GENERAL INFORMATION:
; APPLICANT: Smithson, Glennda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A.
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Ju, Jingfang
; APPLICANT: Li, Li
; APPLICANT: Guo, Xiaojia (Sasha)
; APPLICANT: Patturajan, Meera
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Ellerman, Karen
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Ort, Tatiana
; APPLICANT: Gorman, Linda
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Anderson, David W.
; APPLICANT: Zhong, Mei
; APPLICANT: Catterton, Elina
; APPLICANT: Ji, Weizhen
; APPLICANT: Miller, Charles E.
; APPLICANT: Rastelli, Luca
; APPLICANT: Stone, David J.
; APPLICANT: Pena, Carol E. A.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Leach, Martin D.
; APPLICANT: Agee, Michele L.
; APPLICANT: Berghs, Constance
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-462C
; CURRENT APPLICATION NUMBER: US/10/262,511
; CURRENT FILING DATE: 2003-05-28
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/373,815
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,642
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2002-10-09
; PRIOR APPLICATION NUMBER: 60/381,038
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/373,260
; PRIOR FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 60/373,826
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/327,435
; PRIOR FILING DATE: 2001-10-05
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: CuraSeqList version 0.1
SEQ ID NO 94
; LENGTH: 252
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-511-94

Alignment Scores:
Pred. No.:   3.57e-108   Length: 252
Score: 1252.00   Matches: 228
Percent Similarity: 100.00%   Conservative: 1
Best Local Similarity: 99.56%   Mismatches: 0
Query Match: 70.34%   Indels: 0
DB: 15   Gaps: 0

US-09-905-083A-30 (1-969) x US-10-262-511-94 (1-252)

Qy   88 GCCCAGGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGAGCTCCACCCATGGCAG 147
          ||:::.|||||||||||||||||||||||||||||||||||||||||||||
Db   24 AlaArgGlyAspLysIleIleIleAspGlyAlaProCysAlaArgGlySerHisProTrpGln 43

Qy   148 GTGGCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTGGTCAATGAGCGCTGG 207
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   44 ValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuValAsnGluArgTrp 63

Qy   208 GTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACG 267
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   64 ValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeuGlySerAspThr 83

Qy   268 CTGGGCGACAGGAGGCTCAGAGGACTCAAGGCCTCGAAGTCATTCCGCCACCCGGCTAC 327
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84 LeuGlyAspArgArgLeuArgGlyLeuLysAlaSerLysSerPheArgHisProGlyTyr 103

Qy   328 TCCACAGACCCATGTTAATGACCTCATGTCTCGTGAAGCTCAATAGCCAGGCCAGGCTG 387
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   104 SerThrGlnThrHisValAsnAspLeuMetLeuValLysLeuAsnSerGlnAlaArgLeu 123

Qy   388 TCATCCATGGTGAAGAAAGTCAGGCTGCCTCCGCTGCGAACCCCTGGAACCACCTGT 447
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   124 SerSerMetValLysLysValLysArgLeuProSerArgCysGluProProGlyThrThrCys 143

Qy   448 ACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTCCCTCGACCTCATG 507
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   144 ThrValSerGlyTrpGlyThrThrThrSerProAspValThrPheProSerAspLeuMet 163

Qy   508 TGGGTGGATGTCAAGCTCATCTCCCCAGGACTGCACGAAGGTTACAGGACTTACTG 567
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   164 CysValAspValLysLeuIleSerProGlnAspCysThrLysGlyThrLysTyrLysAspLeuLeu 183

Qy   568 GAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAACGCCTGCAATGGTGAC 627
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   184 GluAsnSerMetLeuCysAlaGlyIleProThrProLysLysAsnAlaCysAsnGlyAsp 203

Qy   628 TCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCCTGGGAACTTTC 687
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   204 SerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSerTrpGlyThrPhe 223

```
Qy   688 CCTTGCGGCCAACCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTCACCAAGTGG 747
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   224 ProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLysPheThrLysTrp 243

Qy   748 ATAAATGACACCATGAAAAGCATCGC 774
          ||||||||||||||||||||||||||
Db   244 IleAsnAspThrMetLysLysHisArg 252
```

Search completed: July 14, 2005, 17:57:24
Job time : 178 secs

```
                GenCore version 5.1.6
        Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on:         July 14, 2005, 17:15:20 ; Search time 45.5 Seconds
                                (without alignments)
                                4098.200 Million cell updates/sec

Title:          US-09-905-083A-30
Perfect score:  1780
Sequence:       1 ggattccggcgtccatggc.........aagaaacacaaacccctcag 969

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext   0.5
                Ygapop 10.0 , Ygapext   0.5
                Fgapop  6.0 , Fgapext   7.0
                Delop   6.0 , Delext    7.0

Searched:       283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters:        566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      PIR_79:*
                1: pirl:*
                2: pir2:*
                3: pir3:*
                4: pir4:*

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09905083/runat_14072005_140353_10622/app_query.fasta_1.1159
-DB=PIR_QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09905083 @CGN_1_63 @runat_14072005_140353_10622 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1364	76.6	253	2	A53968	serine proteinase
2	614	34.5	260	2	I56559	neuropsin - mouse
3	582	32.7	261	2	A29745	tissue kallikrein
4	568	31.9	261	2	S01971	tissue kallikrein
5	563	31.6	261	1	NGMSG	7S nerve growth fa
6	563	31.6	261	2	A34079	tissue kallikrein
7	561	31.5	246	1	TRRT1	trypsin (EC 3.4.21
8	552	31.0	261	1	KQMS1	tissue kallikrein
9	552	31.0	261	2	A31136	tissue kallikrein
10	551	31.0	261	2	A25606	tissue kallikrein
11	547.5	30.8	260	2	A37938	tissue kallikrein
12	541.5	30.4	246	1	TRRT2	trypsin (EC 3.4.21
13	536	30.1	248	2	S55067	trypsin (EC 3.4.21
14	535	30.1	229	1	TRBOTR	trypsin (EC 3.4.21
15	535	30.1	265	1	KQRTP	tissue kallikrein
16	534	30.0	261	1	TRMSM5	tissue kallikrein
17	533.5	30.0	244	2	A44284	tissue kallikrein
18	533.5	30.0	247	2	A27547	trypsin (EC 3.4.21.-
19	533	29.9	246	2	B25528	trypsin (EC 3.4.21
20	533	29.9	259	1	KQRTTN	tonin (EC 3.4.21.-
21	533	29.9	259	2	A29746	tissue kallikrein
22	533	29.9	259	2	B31116	tissue kallikrein
23	532.5	29.9	231	2	S31778	trypsin (EC 3.4.21
24	532	29.9	242	2	S31776	trypsin (EC 3.4.21
25	532	29.9	242	2	S31775	trypsin (EC 3.4.21
26	531	29.8	261	2	A41020	tissue kallikrein
27	527.5	29.6	247	1	TRDG	trypsin (EC 3.4.21
28	527	29.6	261	2	JE0236	tissue kallikrein
29	526.5	29.6	242	2	S49489	trypsin (EC 3.4.21
30	526.5	29.6	247	2	S13813	trypsin (EC 3.4.21
31	524	29.4	231	1	TRPGTR	trypsin (EC 3.4.21
32	521	29.3	261	2	A24378	tissue kallikrein
33	519	29.2	261	1	EGMSB	trypsin (EC 3.4.21
34	518.5	29.2	243	2	A35871	trypsin (EC 3.4.21
35	518.5	29.1	262	1	KQHU	tissue kallikrein
36	516.5	29.0	256	1	NGMSA	7S nerve growth fa
37	516	29.0	257	2	S33772	tissue kallikrein
38	514	28.9	232	1	KQPG	tissue kallikrein
39	513	28.7	261	2	A29586	tissue kallikrein
40	511	28.7	238	2	S31779	trypsin (EC 3.4.21
41	510.5	28.7	247	1	A25852	trypsin (EC 3.4.21
42	509	28.6	263	2	S15686	tissue kallikrein
43	508	28.5	246	1	TRDGC	trypsin (EC 3.4.21
44	506	28.4	261	2	S45303	tissue kallikrein
45	503	28.3	248	2	S55066	trypsin (EC 3.4.21

ALIGNMENTS

RESULT 1

A53968

serine proteinase SCCE precursor - human
N;Alternate names: stratum corneum chymotryptic enzyme
C;Species: Homo sapiens (man)
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 09-Jul-2004
C;Accession: A53968
R;Hansson, L.; Stroemqvist, M.; Baeckman, A.; Wallbrandt, P.; Carlstein, A.; Egelrud, T.
J. Biol. Chem. 269, 19420-19426, 1994
A;Title: Cloning, expression, and characterization of stratum corneum chymotryptic enzym
A;Reference number: A53968; MUID:94308225; PMID:8034709
A;Accession: A53968
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-253 <HAN>
A;Cross-references: UNIPROT:P49862; GB:L33404; NID:g521214; NID:g532504
A;Cross-references: UNIPROT:P49862; GB:L33404; NID:g521214; PIDN:AAC37551.1; PID:g532504
C;Genetics:
A;Gene: GDB:PRSS6; SCCE
A;Cross-references: GDB:377730
A;Map position: 7q35-7q35
C;Superfamily: trypsin; trypsin homology <TRY>
F;30-245/Domain: trypsin homology <TRY>

```
Alignment Scores:
Pred. No.:                                  253
Score:                  3.28e-109            253
Percent Similarity:     1364.00
Best Local Similarity:  100.00%        Length:              253
Query Match:            100.00%        Matches:             253
DB:                     76.63%         Conservative:          0
                        2              Mismatches:            0
                                       Indels:                0
                                       Gaps:                  0

US-09-905-083A-30 (1-969) x A53968 (1-253)

Qy    16 ATGGCAAGATCCCTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT  75
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr  20

Qy    76 GCAGGAGGAGAAGGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCC 135
```

```
US-09-905-083A-30 (1-969) x I56559 (1-260)

Qy      37 CCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAGAAGAGCCAGGGT 96
           ||| ||||||||||| :::||||||||
Db      11 ProTrpIleLeuLeuLeuPheMetGlyAlaTrpAlaGlyLeuThrArgAlaGlnGly 30

Qy      97 GACAAGATTATTGATGGCGCCCATGCCCCAGAGAGGCTCCAAGAGAGCTCCACCATGGCAGGTGGCCCTG 156
           |||||| ||||
Db      31 SerLysIleLeuGlyLysGluArgGluCysIleProlleProHisSerGlnProTrpGlnAlaAlaLeu 50

Qy     157 CTCAGTGGCAATCAGCTCCACTGCGGAGGCCGTTCTGGTCAATGACGCTGGGTGGCTCACT 216
           ||||||||||||||||||||                    :::||||||||
Db      51 PheGlnGlyGluArgLeuIleCysGlyGlyValLeuValGlyAspArgTrpValLeuThr 70

Qy     217 GCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACGTCG----- 270
           ||||||||||||||||||||||||||| :::||||||||
Db      71 AlaAlaHisCysLysLysGlnLysLysTyrSerValArgLeuGlyAspHisSerLeuGlnSer 90

Qy     271 GGGCAGAGGAGAGCTCAGAGGATCAAGGCCTGAAGTCATTCCGCCACCCGGCTACTCC 330
           ||||| ||||
Db      91 ArgAspGlnProGlnGluGlulleGlnValAlaGlnSerlleGlnHisProCysTyrAsn 110

Qy     331 ACA---------CAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGGCC 381
           :::                   ||||||||||
Db     111 AsnSerAsnProGluAspHisSerHisAspIleMetLeuIleArgLeuGlnAsnSerAla 130

Qy     382 AGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCCGTGCGAACCCCTGGAACC 441
           ||| :::||||
Db     131 AsnLeuGlyAspLysValTyrValGlnLeuAsnLeuAlaAsnLeuCysProLysValGlyGln 150

Qy     442 ACCTGTACTGTCTCCGGCTGGGCCACTACCACCAGAGCCAGAGTGTGACCTTTCCCTCTGAC 501
           :::||||||||||
Db     151 LysCysIlelleSerGlyTrpGlyThrValThrSerProGlnGluAsnProAsnThr 170

Qy     502 CTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTACAAGGAC 561
           ||| :::||||||||
Db     171 LeuAsnCysAlaGluValCysAspLysTyrSerGlnAsnLysCysGluArgAlaTyrProGly 190

Qy     562 TTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCCAGAGAAAAAGCCGCTGCAAT 621
           ::: ||||
Db     191 LysIleThrGluGlyMetValCysAlaGly--SerSerAsnGlyAlaAspThrCysGln 209

Qy     622 GGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTCGTGTCCTGGGGA 681
           |||||||||| :::||||
Db     210 GlyAspSerGlyGlyProLeuValCysAspGlyMetLeuGlnGlyIleThrSerTrpGly 229

Qy     682 ACTTTCCCTTGCGGCCAACCCAATGACCCAGGAGTCTACACTCAAGTGTCAAGTTCACC 741
           ::: ||||
Db     230 SerAspProCysGlyGlyProGluLysProGlyValTyrThrLysCysArgTyrThr 249

Qy     742 AAGTGGATAAATGACACCATG 762
           |||||
Db     250 ThrTrpIleLysLysThrMet 256
```

A;Reference number: A29745; MUID:88107593; PMID:3322386
A;Accession: A29745
A;Molecule type: mRNA
A;Residues: 1-261 <BLA>
A;Cross-references: GB:M17962; NID:g192997; PIDN:AAA37541.1; PID:g309212
R;Isackson, P.J.; Silverman, R.E.; Blaber, M.; Server, A.C.; Nichols, R.A.; Shooter, E.M
Biochemistry 26, 2082-2085, 1987
A;Title: Epidermal growth factor binding protein: identification of a different protein.
A;Reference number: A27120; MUID:87299636; PMID:3304419
A;Accession: A27120
A;Molecule type: protein
A;Residues: 25-54;112-124,'X',126-130;165-184,'X',186-187,'X',189-192 <ISA>
R;Evans, B.A.; Drinkwater, C.C.; Richards, R.I.
J. Biol. Chem. 262, 8027-8034, 1987
A;Title: Mouse glandular kallikrein genes: Structure and partial sequence analysis of th
A;Reference number: I55260; MUID:87250386; PMID:3036794
A;Accession: I70015
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 70-122 <RES>
A;Cross-references: GB:M18608; NID:g198500; PIDN:AAA39351.1; PID:g198506
C;Comment: This sequence is one of approximately twenty-five members of a gene family on
C;Genetics:
A;Gene: KAL
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase; zymogen
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: propeptide #status predicted <PRO>
F;25-261/Product: tissue kallikrein mGK-9, submandibular #status experimental <MAT>
F;25-253/Domain: trypsin homology <TRY>
F;65,120,213/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.:   Length: 261
Score: 5.92e-42   Matches: 115
Percent Similarity: 59.23%   Conservative: 39
Best Local Similarity: 44.23%   Mismatches: 90
Query Match: 32.70%   Indels: 16
DB: Gaps: 3

US-09-905-083A-30 (1-969) x A29745 (1-261)

Qy   40 CTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGGAGAAGAGCCCAG------ 93
            :::||| |||:|||| |||||||| || ::||||
Db   1 MetArgPheLeuIleLeuPheLeuAlaLeuSerLeuGlyGlyIleIleAspAlaAlaProPro 20

Qy   94 ---GGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCCACCATGGCAGGTG 150
            ::||||::: ||| ::||| |||||
Db   21 ValHisSerArgIleValGlyPheLysCysGluLysSerGlnProTrpHisVal 40

Qy   151 GCCCTGCTCAGTGGCAATCAGCTCCACTGCGAGGCGTTCCTGGTCAATGAGCGCTGGGTG 210
            |||:::: ||||:|||||||||||| ||:::::
Db   41 AlaValTyrArgTyrAsnGluTyrIleCysGlyGlyValLeuLeuAspAlaAsnTrpVal 60

Qy   211 CTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACGCTG 270
            |||||||||| ::|| :: |||||
Db   61 LeuThrAlaAlaHisCysTyrTyrTyrGluGluSerLeuGlyLysValSerLeuLysAsnAsnLeu 80

Qy   271 GGGCAGCAGGAGA------GCTCAGAGGATCAAGGCCTGAAGTCATTCCGCACCCCCGGC 324
            ::: ||||||| |||||||||||
Db   81 TyrGluGluProSerAlaGlnHisArgLeuValSerLysSerPheLeuHisProGly 100

Qy   325 TACTCCACACAGACCCAT-----------------------------GTTAATGAC 351
            |||:::||| |||||||
Db   101 TyrAsnArgSerLeuHisArgAsnHisIleArgHisProGluTyrSerProGluTyrSerProTyrSerAsnAsp 120

Qy   352 CTCATGCTCGTGAAGCTCAATAGCCAGGCCAGGTGTCATCCATGGTGAAGAAGTCAGG 411
            |||::::||| |||||:||::||||||||
Db   121 LeuMetLeuValLysLeuAsnSerGlnAlaArgGlnProAlaAspAspAlaProValIleProIleAla 140

Qy   412 CTGCCCTCCCGGTGACAACCCCCTGGAACCACCTTACTGTCTCCCCGCTGGGGCCACTACC 471
            ||||::::||| |:||:||||||||||||||||||||||||:
Db   141 LeuProThrGluGluProLysProLeuLeuGlySerThrCysLeuAlaSerGlyTrpGlySerThr 160

RESULT 4
S01971
tissue kallikrein (EC 3.4.21.35) precursor - mouse
N;Alternate names: glandular kallikrein
C;Species: Mus musculus (house mouse)
C;Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 09-Jul-2004
C;Accession: S01971; I70023
R;Drinkwater, C.C.; Richards, R.I.
Nucleic Acids Res. 16, 10918, 1988
A;Title: Sequence of mGK-11, a mouse glandular kallikrein gene.
A;Reference number: S01971; MUID:89083511; PMID:3205728
A;Accession: S01971
A;Molecule type: DNA
A;Residues: 1-261 <DRI>
A;Cross-references: UNIPROT:P15946; EMBL:X13215; NID:g53064; PIDN:CAA31604.3; PID:g57082
R;Evans, B.A.; Drinkwater, C.C.; Richards, R.I.
J. Biol. Chem. 262, 8027-8034, 1987
A;Title: Mouse glandular kallikrein genes: Structure and partial sequence analysis of th
A;Reference number: I55260; MUID:87250386; PMID:3036794
A;Accession: I70023
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 70-122 <RES>
A;Cross-references: GB:M18610; NID:g198529; PIDN:AAA39353.1; PID:g198537
C;Genetics:
A;Introns: 16/1; 69/2; 165/1; 210/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-24/Domain: propeptide #status predicted <PRO>
F;25-261/Product: tissue kallikrein #status predicted <MAT>
F;25-253/Domain: trypsin homology <TRY>
F;65,120,213/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.:   Length: 261
Score: 9.47e-41   Matches: 112
Percent Similarity: 57.98%   Conservative: 37
Best Local Similarity: 43.58%   Mismatches: 92
Query Match: 31.91%   Indels: 16
DB: Gaps: 3

US-09-905-083A-30 (1-969) x S01971 (1-261)

Qy   49 TTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGGAGAAGAGCCCAG--------GGTGAC 99
            |||:::||| |||||||| || ::|||
Db   4 LeuIleLeuPheLeuAlaLeuSerLeuGlyGlyIleIleAspAlaAlaProValGlnSer 23

Qy   100 AAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCCACCATGGCAGGTGGCCCTGCTC 159
            :::|||| || ||| ::||| |||||
Db   24 ArgIleValGlyPheAsnCysGluLysAsnSerGlnProTrpHisValAlaValTyr 43

---

472 ACGAGCCCAGATGTGACCTTCCCTGACCTGGATGCTGGAGCTCAAGCTCATCTCC 531
    ||| ||||||||||| ||||||| :::||||||:::
161 ThrProLysPheGlnAsnAlaAlaLysAspLeuGlnCysValAsnLysLeuLysLeuPro 180

532 CCCCAGGACTGCACGAAGGTTTACAAGGACTTACTGGAAATTCCATGCTGTGGCGGCTGGC 591
    :::|||||| || ::: ::| :::   ::|||
181 AsnGluAspCysGlyLysLysAlaHisIleGluLysValThrAspValMetLeuCysAlaGly 200

592 ATCCCCGACTCCAAGAAAAACGCCTGCAATGGACAAGTGACTCAGGGGGACCGTTGGTGTCGAGA 651
    |||:|| ||| |||||||||||||||||||||||||||
201 GluThrAspGlyGlyLysAspThrCysLysGlySerGlyGlyProLeuIleCysAsnGlyAsp 220

652 GGTACCCTGCAAGGTCTGGTGTCCTGGGGAACTTTCCCTTGCGGCCAACCCAATGCACCA 711
    |||||||::: |||||||||||| ||||
221 GlyValLeuGlnGlyIleThrSerTrpGlyPheThrProCysGlyGluProLysLysPro 240

712 GGAGTCTACACTCAAGTGCCAAGTTCACCCAAGTGGATAAATGACACCATGAAAAAGCAT 771
    |||||||||| |||::::: ||||||| :::|||::|
241 GlyValTyrThrLysLeuTyrLysLysPheThrSerThrTrpIleLysAspThrMetAlaLysAsn 260

A;Cross-references: GB:X00472; NID:g54260; PIDN:CAA25154.1; PID:g54261
A;Experimental source: inbred strain DBA/2J
R;Thomas, K.A.; Baglan, N.C.; Bradshaw, R.A.
J. Biol. Chem. 256, 9156-9166, 1981
A;Title: The amino acid sequence of the gamma-subunit of mouse submaxillary gland 7 S ne
A;Reference number: A92341; MUID:81264363; PMID:7263706
A;Accession: A92341
A;Molecule type: protein
A;Residues: 25-107,112-261 <THO>
A;Experimental source: outbred strain Swiss Webster
C;Comment: 7S nerve growth factor is composed of two alpha chains, a beta dimer composed
C;Comment: The active form of the gamma chain occurs naturally as combinations of either
C;Genetics:
A;Map position: 7
A;Introns: 16/1; 69/2; 165/1, 210/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; growth factor; hydrolase; serine proteinase; submandibular gla
F;1-18/Domain: signal sequence #status predicted <SIG>
F;25-253/Domain: trypsin homology <TRY>
F;25-107/Domain: segment B1 <GB1>
F;25-107,112-261/Product: nerve growth factor gamma chain (active form) #status experime
F;112-261/Domain: segment A <GAA>
F;112-164/Domain: segment C <GCC>
F;165-261/Domain: segment B2 <GB2>
F;31-173,50-66,152-219,184-198,209-234/Disulfide bonds: #status predicted
F;65,120,213/Active site: His, Asp, Ser #status predicted
F;102/Binding site: carbohydrate (Asn) (covalent) #status experimental

Alignment Scores:
Pred. NO.:
Score: 2.55e-40   Length: 261
                                                    Matches: 112
                                                    Conservative: 42
Percent Similarity: 59.92%   Mismatches: 87
Best Local Similarity: 43.58%   Indels: 16
Query Match: 31.63%   Gaps: 4
DB: 1

US-09-905-083A-30 (1-969) x NGMSG (1-261)

Qy   49 TTACTGCTATCCTTAGCCTGGAAACTGCAGGAGAGAAGAGCCAG--------GGTGAC 99
            ||:||||| ||| |||||||||||||||||||||| :::|||
Db   4 LeuLeuLeuPheLeuAlaLeuSerLeuGlyGlyIleAspAlaAlaProProValGlnSer 23

Qy 100 AAGATTATTGATGGCGCCCATGTGCAAGAGAGGCTCCCACCCATGCAGTGGCCCTGCTC 159
            :::||||:::| ||| |||::   ::: |||||||||||||||||||||
Db   24 ArgIleValGlyPheLysCysGluLysAsnSerGlnProTrpHisValAlaValTyr 43

Qy 160 AGTGGCAATCAGCTCCACTGCGCGGAGGCGTCCTGGTCAATGAGCCGGTGCTCCACTGCC 219
            ||||||||||||:::||||||||||||||||||||||||
Db   44 ArgTyrThrGlnTyrLeuCysGlyGlyValLeuLeuAspArgAsnTrpValLeuThrAla 63

Qy 220 GCCCACTGCGAAGATGAATGAGACACCGTGCCACCTGGGCAGTGATACCGTG------GGC 273
            |||||||||||| ||| :::: ||||||||||||||||||||||| |||
Db   64 AlaHisCysTyrAspAspAsnTyrLysValTrpLeuGlyLysAsnAsnLeuPheLysAsp 83

Qy 274 GACAGGAGAGCTCAGGATGACAAGGCCTCGAAGTCATTCCGACCACCCGGCTACTCC--- 330
            ||||||||:::||||||||||||||||| ::::|||||||||||||||||||
Db   84 GluProSerAlaGlnHisArgPheValSerLysAlaIleProHisProGlyPheAsnMet 103

Qy 331 ------------------AATGACCTCATGCTC 360
            :: :::::::||||||||||||||||
Db 104 SerLeuMetArgLysPheHisIleArgPheLeuGluTyrSerAsnAspLeuMetLeu 123

Qy 361 GTGAAGCTCAATAGCCAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGTGCCCTCC 420
            :::::||||||||:::||||||||||||||||||||||||||||||||||||||
Db 124 LeuArgLeuSerLysProAlaAspIleThrAspAlaValLysProIleAlaLeuProThr 143

Qy 421 CGCTGCGACCCCCTGGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCA 480
            ||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db 144 LeuArgLeuSerGluThrLysProAlaAspIleAspAlaValLysProIleAlaLeuProThr
            (continuing)

Qy 481 GATGTGACCTTTCCCTCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGAC 540
            ||| :::|||||||:::|||||||||||| ||| :::|||
Db 164 LysPheGlnThrProAspAspLeuProAsnLeuLysLeuProAsnGluVal 183

Qy 541 TGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCCGAC 600
            |||||||||||||||::::||||||||||||||||||||||||||||||||||
Db 184 CysValLysAsnHisAsnGlnLysValThrAspValMetLeuCysAlaGlyIleGluMetGly 203

Qy 601 TCCAAGAAAACGCCTGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTG 660
            |||||||||| :::||||||||||||||::::||||||||||||||||||||||||
Db 204 GlyGlyLysAspThrCysGlnGlyAspSerGlyGlyProLeuIleCysAspGlyValLeu 223

Qy 661 CAAGGTCTGGTGTCCTGGGGAACTTCCCTGCGGGCAACCCAATGACCAGGAGTCTAC 720
            |||:::||| ::|||||||||||:::||||||||||||||:::
Db 224 HisGlyIleThrAlaTrpProIleProGlyProCysGlyLysProGlyAsnThrProGlyValTyr 243

Qy 721 ACTCAAGTGTGCAAGTTCACCAAGTTCGGATAAAATGACACCATGAAAAAAGCAT 771
            ||:|||:::||||||||||::||||||||| ||| :::|||
Db 244 ThrLysLeuIleLysPheThrAsnTrpIleLysAspThrMetAlaLysAsn 260

RESULT 5
NGMSG
7S nerve growth factor gamma chain (EC 3.4.21.-) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 18-Dec-1981 #sequence_revision 17-May-1985 #text_change 09-Jul-2004
A;Accession: A91005; A90949; A93510; A92341; A00942; A21093; A22705
R;Evans, B.A.; Richards, R.I.
EMBO J. 4, 133-138, 1985
A;Title: Genes for the alpha and gamma subunits of mouse nerve growth factor are contigu
A;Reference number: A91005; MUID:85257431; PMID:3848399
A;Accession: A91005
A;Molecule type: DNA
A;Residues: 1-261 <EVA>
A;Cross-references: UNIPROT:P00756
R;Ullrich, A.; Gray, A.; Wood, W.I.; Hayflick, J.; Seeburg, P.H.
DNA 3, 387-392, 1984
A;Title: Isolation of a cDNA clone coding for the gamma-subunit of mouse nerve growth fa
A;Reference number: A90949; MUID:85076169; PMID:6548955
A;Accession: A90949
A;Molecule type: mRNA
A;Residues: 1-261 <ULI>
A;Cross-references: GB:X01389; NID:g53373; PIDN:CAA25645.1; PID:g53374
R;Howles, P.N.; Dickinson, D.P.; DiCaprio, L.L.; Woodworth-Gutai, M.; Gross, K.W.
Nucleic Acids Res. 12, 2791-2805, 1984
A;Title: Use of a cDNA recombinant for the gamma-subunit of mouse nerve growth factor to
A;Reference number: A93510; MUID:84169573; PMID:6200835
A;Accession: A93510
A;Molecule type: mRNA
A;Residues: 127-202,'E',204-261 <HOW>

Qy   541 TGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATGTGCGCTGGCATCCCCGAC 600
          ||| :::   :::   ::: ||||| ::: ||
Db   184 CysAlaLeuAlaHisIleGluLeuValThrAspAlaMetLeuCysAlaGlyGluMetAsp 203

Qy   601 TCCAAGAAAAACGCCTGCAATGCTGGTGACTCAGGGGGACCGTTGGTGTGTGCAGAGGTACCCTG 660
          ||| ::: |||
Db   204 GlyGlyLysAspThrCysLysGlyAspSerGlyGlyProLeuIleCysAspGlyValLeu 223

Qy   661 CAAGGTCTGGTGTCCTGGGGAACTTTCCCTGCGGCCAACCCAATGACCAGGAGTCTAC 720
Db   224 GlnGlyIleThrSerTrpGlyHisThrProCysGlyGluProAspMetProGlyValTyr 243

Qy   721 ACTCAAGTGTGCAAGTTCACCAAGTTGGATAAATGACACACCATGAAAAAAGCAT 771
          |||:::: |||||
Db   244 ThrLysLeuAsnLysPheThrSerTrpIleLysHisAspThrMetAlaLysAsn 260

RESULT 6
A34079
tissue kallikrein (EC 3.4.21.35) P1 precursor - rat
N;Alternate names: kallikrein-related proteinase k8
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Mar-1990 #sequence_revision 30-Mar-1990 #text_change 09-Jul-2004
C;Accession: A34079; S10700
R;Brady, J.M.; Wines, D.R.; MacDonald, R.J.
Biochemistry 28, 5203-5210, 1989
A;Title: Expression of two kallikrein gene family members in the rat prostate.
A;Reference number: A34079; MUID:89352606; PMID:2765531
A;Accession: A34079
A;Status: preliminary
A;Molecule type: DNA; mRNA
A;Residues: 1-261 <BRA>
A;Cross-references: UNIPROT:P36374; GB:M27215; GB:M27216; GB:M27217; NID:g206638; PIDN:A
A;Experimental source: prostate
R;Elmoujahed, A.; Gutman, N.; Brillard, M.; Gauthier, F.
FEBS Lett. 265, 137-140, 1990
A;Title: Substrate specificity of two kallikrein family gene products isolated from the
A;Reference number: S10698; MUID:90306305; PMID:2194829
A;Accession: S10700
A;Molecule type: protein
A;Experimental source: submaxillary gland
A;Note: 125-Lys was also found
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;25-253/Domain: trypsin homology <TRY>
F;25-111/Product: tissue kallikrein P1 light chain #status experimental <MAT1>
F;112-261/Product: tissue kallikrein P1 heavy chain #status experimental <MAT2>
F;65,120,213/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.:
Score: 2.55e-40   Length: 261
Percent Similarity: 563.00   Matches: 112
Best Local Similarity: 58.30%   Conservative: 39
Query Match: 43.24%   Mismatches: 94
DB: 31.63%   Indels: 14
                         2   Gaps: 3

US-09-905-083A-30 (1-969) x A34079 (1-261)

Qy   34   CTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGGAAGAAGCCCAG 93
          ||| ||| | :||:||||| || :|||||||||||
Db   3   LeuLeuIleLeuPheLeuIleLeuLeuSerLeuGlyTrpAsnAspAlaAlaProProGlyGln 22

Qy   94   GGTGACAAGATTAATGATGGCGCCCCATGTGCAGAGAGCTCCCACCCATGGCAGGTGGCC 153
          :: :|||| || ||| ::
Db   23   --SerArgIleGluGlyPheAsnCysGluLysAsnSerGlnProTrpGlnValAla 41

Qy   154 CTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTCGTCAATGAGCGGTGGGTGCTC 213
          |||:|| ||||::|
Db   42   ValTyrHisPheAsnGluProGlnCysGlyGlyValLeuIleHisProSerTrpValIle 61

Qy   214 ACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACGCTG--- 270

Db ||||||||||||
Db   62   ThrAlaAlaHisCysTyrSerValAlaAsnTyrGlnValTrpLeuGlyArgAsnAsnLeuLeu 81

Qy   271 --GGCGACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGCTCATTCCGCCACCCGGCTAC 327
          ::: |||
Db   82   GluAspGluProPheAlaGlnHisArgLeuValSerGlnInSerPheProHisProGlyPhe 101

Qy   328 TCCCA------------------------------CAGACCCATGTTAATGACCTC 354
          :::
Db   102 AsnLeuAspAsnIleLeuLysAsnHisThrArgLysProGlyAsnAspAspTyrSerAsnAspLeu 121

Qy   355 ATGCTCGTGAAGCTCAATAGCCAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTG 414
          ||||| ::: ||
Db   122 MetLeuLeuHisLeuLysThrProAlaAspIleThrAspGlyValLysValIleAspLeu 141

Qy   415 CCCTCCCGCTGGAACACCCCTGGAACCACCTGTCTCTCCGGCTGGGGCACTACCACG 474
          ||||:::: |||||
Db   142 ProThrGluProLysValGlySerThrCysLeuThrSerGlyTrpGlySerIleThr 161

Qy   475 AGCCCAGATGTGACCTTCCCTGCACCTCCATGTGCGATGTGAGTCAAGCTCATCTCCCC 534
          |||:||:|||| ::: |||
Db   162 ProLeuLysTrpGluPheProAspAspLeuGlnCysValAsnIleHisLeuLeuSerAsn 181

Qy   535 CAGGACTGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATGTGCGCTGGCATC 594
          ::: ::: ||| ::: |||
Db   182 GluLysCysIleLysLysAlaTyrAsnAspSerGluValThrAspAlaMetLeuCysAlaGlyGlu 201

Qy   595 CCCGACTCCAAGAAAAACGCCTGCAATGGTGACTCAGGGGGACCGTTGGTGTGTGCAGAGGT 654
          ||||||::: |||
Db   202 MetAspGlyGlyLysAspThrCysLysGlyAspSerGlyGlyProLeuIleCysAspGlyIly 221

Qy   655 ACCCTGCAAGGTCTGGTGTCCTGGGGAACTTTCCCTGCGGCCAACCCAATGACCAGGA 714
          ||||:||||| ::: |||
Db   222 ValLeuGlnGlyIleThrSerTrpGlyMetProCysGlyGluProAsnLysProSer 241

Qy   715 GTCTACACTCAAGTGTGCAAGTTCACCAAGTTGGATAAATGACACACCATGAAAAAGCAT 771
          ||||| ::::: |||
Db   242 ValTyrThrLysLeuLysLysPheThrSerTrpMetLysValMetLysGluAsn 260

RESULT 7
TRRT1
trypsin (EC 3.4.21.4) I precursor - rat
N;Alternate names: trypsinogen I
C;Species: Rattus norvegicus (Norway rat)
C;Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 09-Jul-2004
C;Accession: B22657; A00948
R;Craik, C.S.; Choo, Q.L.; Swift, G.H.; Quinto, C.; MacDonald, R.J.; Rutter, W.J.
J. Biol. Chem. 259, 14255-14264, 1984
A;Title: Structure of two related rat pancreatic trypsin genes.
A;Reference number: A22657; MUID:85054880; PMID:6094547
A;Accession: B22657
A;Molecule type: DNA
A;Residues: 1-246 <CRA>
A;Cross-references: UNIPROT:P00762; GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
A;Note: the authors translated the codon ATC for residue 6 as Leu and GAC for residue 17
R;MacDonald, R.J.; Stary, S.J.; Swift, G.H.
J. Biol. Chem. 257, 9724-9732, 1982
A;Title: Two similar but nonallelic rat pancreatic trypsinogens. Nucleotide sequences of
A;Reference number: A00948; MUID:82265624; PMID:6896710
A;Accession: A00948
A;Molecule type: mRNA
A;Residues: 1-246 <MAC>
A;Cross-references: GB:J00778; NID:g206507; PIDN:AAA98518.1; PID:g206508
A;Genetics:
A;Introns: 14/1; 67/2; 152/1; 197/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-23/Domain: activation peptide #status predicted <APT>
F;24-246/Product: trypsin I #status predicted <ENZ>
F;24-239/Domain: trypsin I homology <TRY>
F;30-160,48-64,132-233,139-206,171-185/Disulfide bonds: #status predicted
F;63,107,200/Active site: His, Asp, Ser #status predicted

F;75,77,80,85/Binding site: calcium (Glu, Asn, Val, Glu) #status predicted

Alignment Scores:
Pred. No.:   246
Score: 3.77e-40   Matches: 113
Percent Similarity: 60.08%   Conservative: 39
Best Local Similarity: 44.66%   Mismatches: 87
Query Match: 31.52%   Indels: 14
DB: Gaps: 5

US-09-905-083A-30 (1-969) x TRRT1 (1-246)

```
Qy     16 ATGGCAAGATCCCTTCTCCTGCCCCTGCCAGATCTTACTGCTATCCTTAGCCTTGGAAACT 75
          ||| ::::||||||  |||::::||  ||:|::| |||::::| | ||   ::::||
Db      1 MetSerAlaLeuLeuIleLeuLeuAlaLeuValGlyAlaAlaLeuValAlaValAlaPheProLeuGlu-- 19

Qy     76 GCAGGAGAAGAAGCCCAGGTGACAAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCC 135
          |||||||||||::|:    :::||||||||      |::| |    |||:||::||||
Db     20 ----------AspAspAspLysIleValGlyGlyTyrThrCysProGluHisSer 34

Qy    136 CACCCATGCCAGGTGGCCCTGCTCAGTGCAATCAGCTCCACTGCGGAGGCGGTCCTGGTC 195
          ||||| |||||||||||:::: ||||     |||  |||  |||:||||||:||||||:::
Db     35 ValProTyrGlnValSerLeuAsnSerGlyTyrHisPhe---CysGlyGlySerLeuIle 53

Qy    196 AATCAGGCGCTGGGTGCTCACTGCGCCCACTGACAAGATGAATGAGTACACCGTGCACCTG 255
          |||| :||||||||||||||:::  |||   ||| |||   ::: ||||  ||:|||||
Db     54 AsnAspGlnTrpValValSerArgIleGlnValArgLeu 73

Qy    256 GGCAGTGATACGCTG----------GGCGACAGGAGAGCTCAGAGGATCAAGGCCTCG 303
          ::: ||||||    ||||||      ||||    ::::|:|| |||||||
Db     74 GlyGluHisAsnIleAsnValGluGlyGluGlu------GlnPheIleAsnAlaAla 91

Qy    304 AAGTCCATTCCGCCACCCGGCTACTCCACACAGACCCATGTTAATGACCTCATGTCGTG 363
          ||| ::::||||||||||||||||||   ::: ||||| |||  |||   ||||:||||
Db     92 LysIleIleLysHisProAsnTyrSerSerTrpThrLeuAsnAspIleMetLeuIle 111

Qy    364 AAGCTCAATAGCCAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCTCCCGGC 423
          |||||||||||||||||||||||||| |||      |::||::| |||  ||||:::
Db    112 LysLeuSerSerProValLysLeuAsnAlaArgValAlaAlaProValAlaLeuProSerAla 131

Qy    424 TGGCAACCCCCTGCAACCACCTGTCTCCGGCTGGGGACTACCACGAGCCCAGAT 483
          |||||||||||||||||||| :::    ||||||||| ||||| |||:::||||
Db    132 CysAlaProAlaGlyThrGlnCysLeuIleSerGlyTrpGlyTyrAsnThrSerAsnGly 151

Qy    484 GTGACCTTTCCCTGCCTGACCCATGTGCGTGGGGATGTCAAGCTCATCTCCCCCAGGACTGC 543
          |||||||||||||||     |||||||:||||||||||  |||||:::  ||||||||
Db    152 ValAsnAsnProAspLeuLeuGlnCysValAspAlaProValIleLeuSerGlnAlaAspCys 171

Qy    544 ACGAAGGTTTACAAGGACTTACTGGAAAATTCCATGCTGTGCCGCTGGCATCCCGACTCC 603
          ||||||||||| :::|||||||||||| |||||     |:||:|||:|    ::: 
Db    172 GluAlaAlaAlaTyrProGlyLeuIleThrSerSerMetIleCysValGlyPheLeuGluGly 191

Qy    604 AAGAAAAACGCCTGCAATGGTGACTCAGGGGACCGTTGGTGTGCAGAGGTACCCTGCAA 663
          |||||||||||||||||||  |::: |||||||||||:::||||:::
Db    192 GlyLysAspSerCysGlnGlyAspSerGlyGlyProValGlyProValCysAsnGlyGlnLeuGln 211

Qy    664 GGTCTGGTGTCCTGGGAACTTTCCCTTGCGGCCCAACCCAATGACCCCAGGAGTCTCACACT 723
          |||||||||:::     :::::||||  |::: ||||||:::|||||||||||
Db    212 GlyIleValSerTrpGly---TyrGlyCysAlaLeuProAspAsnProGlyValTyrThr 230

Qy    724 CAAGTGTGCAAGTTCACCAAGTGGATAAATGACACCATG 762
          ::::||||||||||| ||||||||||||||:::
Db    231 LysValCysAsnPheValGlyGlnAspTrpIleGlnAspThrIle 243
```

RESULT 8

KQMS1
tissue kallikrein (EC 3.4.21.35) mGK-1 precursor, submandibular - mouse
N;Alternate names: glandular kallikrein; kininogenin
C;Species: Mus musculus (house mouse)
C;Date: 19-Feb-1984 #sequence_revision 19-Feb-1984 #text_change 09-Jul-2004
C;Accession: A00941
R;Mason, A.J.; Evans, B.A.; Cox, D.R.; Shine, J.; Richards, R.I.
Nature 303, 300-307, 1983

A;Title: Structure of mouse kallikrein gene family suggests a role in specific processin
A;Reference number: A00941; MUID:83219214; PMID:6602295
A;Accession: A00941
A;Molecule type: DNA
A;Residues: 1-261 <MAS>
A;Cross-references: UNIPROT:P00755; GB:V00829; NID:g52775; GB:V00755; PIDN:CAA24213.1; PID:g52776
A;Experimental source: Quakenbush inbred strain
C;Comment: mGK-1 belongs to a family of 25 to 30 homologous kallikrein genes.
C;Comment: Tissue kallikreins cleave Met-Lys and Arg-Ser bonds in kininogen to release 1
C;Genetics:
A;Gene: mGK-1
A;Map position: 7
A;Introns: 16/1, 69/2; 165/1; 210/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; saliva; serine proteinase; submandibular gland; zym
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-24/Domain: activation peptide #status predicted <APT>
F;25-261/Product: tissue kallikrein, submandibular #status predicted <MPT>
F;25-253/Domain: trypsin homology <TRY>
F;31-173,50-66,152-219,184-198,209-234/Disulfide bonds: #status predicted
F;65,120,213/Active site: His, Asp, Ser #status predicted
F;102/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.:   261
Score: 2.25e-39   Matches: 109
Percent Similarity: 57.59%   Conservative: 39
Best Local Similarity: 42.41%   Mismatches: 93
Query Match: 31.01%   Indels: 16
DB: Gaps: 3

US-09-905-083A-30 (1-261) x KQMS1 (1-261)

```
Qy     49 TTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAAGCCCAG----------GGTGAC 99
          ||||||||||||  ||||||||||||||||||||||||||||         |||
Db      4 LeuLeuLeuPheLeuAlaLeuSerLeuGlyGlyIleAspAlaAlaProValGlnSer 23

Qy    100 AAGATTATTGATGGCGCCCCATGTGCAAGAGGCTCCCACCCATGCCAGGTGGCCCTGCTC 159
          :::||  | |::: |||  |||:||::|| |||   ::: ||||||:::||  ||||||
Db     24 ArgIleValGlyPheLysCysGluLysSerGlnProTrpProHisAlaAlaValTyr 43

Qy    160 AGTGGCAATCAGCTCCACTGCGGAGGCGGTCCTGGTCAATGAGCCTGGTGGGGTCCTCACTGCC 219
          |||||||||  :::  |||:||||||:|||||||||||| ::::|| ||||::::
Db     44 ArgTyrLysGluTyrIleCysGlyGlyValGlyValLeuLeuLeuAspArgAlaAsnTrpLeuThrAla 63

Qy    220 GCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGCAGTGATACCCTG------GGC 273
          |||||||||||||||| ::: |||||:::||||||||||:::||||| ||      :::
Db     64 AlaHisCysTyrTyrGluLysAsnAsnValTrpLeuGlyLysAsnAsnLeuTyrGlnAsp 83

Qy    274 GACAGGAGAGCTCAGAGGATCAAGGCTCAGAGTCATCCGCCACCCGGCTACTCCACA 333
          |||||||||  ||||| ||| |||   :::  ||:|||  :::|||  ||||||||::
Db     84 GluProSerAlaGlnHisArgGlnValSerPheLeuHisProCysTyrAsnMet 103

Qy    334 CAGACCCATGTTAAT----------------------GACCTCATGTCTC 360
          ||| :::|||  |||                      |||||||||:::
Db    104 SerLeuHisArgAsnArgIleGlnAsnProAspAspTyrSerLeuMetLeu 123

Qy    361 GTGAAGCTCAATAGCCAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCTCC 420
          ||||| |||||||||||||||||||||| |||   |:::||||::| |||  ::: ::: 
Db    124 LeuArgLeuSerProAlaAspIleThrAspValValLysProIleAlaLeuProThr 143

Qy    421 CGGCTGCGAACCCCTGGAACCACCTGTCTCCGGTGGGCACTACCACGAGCCCCA 480
          ::|||||| |||||| |||||||||| ::: ||||||||  |||| |||:::
Db    144 GluGluProLysLeuGlySerThrCysLeuAlaSerGlyTrpGlySerIleIleGluProVal 163

Qy    481 GATGTGACCTTTCCCTGCACCATGTGCGTGATGTCAAGCTCATCTCCCCCAGGAC 540
          |::: |||||||||:::||||| |||  ::: ||:||||||| ::::||
Db    164 LysPheGlnTyrAlaLysAspLeuGlnCysValAsnLeuLysLeuProAsnGluAspAsp 183

Qy    541 TGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATGCTGTGCGCGTGCATCCCGAC 600
          |||  |  ::::||||||| :::     ||||||::||:|||::|:::||||||||::
Db    184 CysAspLysAlaTyrValGlnLysValThrAspValMetLeuCysAlaGlyValLysLysGly 203
```

```
Qy   601 TCCAAGAAAAACGGCTGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTG 660
             ||||||   |||   ||||:||||||||||||::||:|| ||||  |||||||||||
Db   204 GlyGlyLysAspThrCysLysGlyAspSerGlyGlyProLeuIleCysAspGlyValLeu 223

Qy   661 CAAGGTCTCGTGTCTCGGGGAACTTCCCTGCGCCAACCCAGTGACCCAGGAGTCTAC 720
             ||||||||||  ||||||||||  |||||:::||||  ||||||||||||||
Db   224 GlnGlyLeuThrSerTrpGlyTyrAsnProCysGlyProLysGluProLysGlyValTyr 243

Qy   721 ACTCCAAGTGTGCAAGTTCACCAAGTGGATAAAATGACACCATGAAAAAGCAT 771
             ||||:::|| |||||||||||||||| ||||||||  :::::||
Db   244 ThrLysLeuIleLysPheThrSerTrpIleLysAspThrLeuAlaGlnAsn 260

RESULT 9
A31136
tissue kallikrein (EC 3.4.21.35) 7 precursor, submandibular - rat
N;Alternate names: glandular prokallikrein 7, submandibular; proteinase A
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: A31136; S10698; S10699; D41429; B41429; S09315
R;Chen, Y.P.; Chao, J.; Chao, L.
Biochemistry 27, 7189-7196, 1988
A;Title: Molecular cloning and characterization of two rat renal kallikrein genes.
A;Reference number: A31136; MUID:89088074; PMID:2849988
A;Accession: A31136
A;Molecule type: DNA
A;Residues: 1-261 <CHE>
A;Cross-references: UNIPROT:P36373; GB:M19647; GB:J02837; NID:g204999; PIDN:AAA41461.1;
R;Elmoujahed, A.; Gutman, N.; Brillard, M.; Gauthier, F.
FEBS Lett. 265, 137-140, 1990
A;Title: Substrate specificity of two kallikrein family gene products isolated from the
A;Reference number: S10698; MUID:90306305; PMID:2194829
A;Accession: S10698
A;Molecule type: protein
A;Residues: 25-36 <BLM>
A;Accession: S10699
A;Molecule type: protein
A;Residues: 112-139 <EL2>
R;Kato, H.; Nakanishi, E.; Enjyoji, K.; Hayashi, I.; Oh-Ishi, S.; Iwanaga, S.
J. Biochem. 102, 1389-1404, 1987
A;Title: Characterization of serine proteinases isolated from rat submaxillary gland: wi
A;Reference number: A41429; MUID:88198057; PMID:3482210
A;Accession: D41429
A;Molecule type: protein
A;Residues: 112-133 <KAT>
A;Accession: B41429
A;Molecule type: protein
A;Residues: 25-34,'D',36-45,'S',47-67,'X',69-75 <KA2>
R;Brady, J.M.; MacDonald, R.J.
Arch. Biochem. Biophys. 278, 342-349, 1990
A;Title: The expression of two kallikrein gene family members in the rat kidney.
A;Reference number: S09315; MUID:90225801; PMID:2183721
A;Accession: S09315
A;Molecule type: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 43-45,'S',47-114,'A',116-261 <BRA>
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; serine proteinase
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-261/Product: tissue prokallikrein 7, submandibular #status predicted <MAT>
F;25-253/Domain: trypsin homology <TRY>
F;65,120,213/Active site: His, Asp, Ser #status predicted
```

Alignment Scores:

Pred. No.:			
Score:	2.25e-39	Length:	261
Percent Similarity:	552.00	Matches:	110
Best Local Similarity:	57.81%	Conservative:	38
Query Match:	42.97%	Mismatches:	94
DB:	31.01%	Indels:	14
	2	Gaps:	3

```
US-09-905-083A-30 (1-969) x A31136 (1-261)

Qy    46 ATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAGAAGAGCCCAGGGT---GACAAG 102
```

```
Db       ||||||   |||   ||||::|||
Qy     5 IleLeuPheLeuAspLeuSerLeuGlyGlnIleAspAlaAlaProProGlyGlnSerArg 24

Qy   103 ATTATTGATGGCGCCCATGCAAGAGGTCCACCATGGCAGTGGCCCTGCTCAGT 162
             :::||||||   ||||||   |||||   ||||||||||||||||||
Db    25 ValIleGlyTyrTyrLysCysGluLysAsnSerGlnProTrpGlnValAlaLeuTyrSer 44

Qy   163 GGCAATCAGCTCCACTGGCGAGAGGGTCCTGGTCAATGAGCGGCTGGTGCTACTGCCGCC 222
             :::  ||||||||||||||||||||||||||:::||:||||||||||||||:::|||
Db    45 PheThrLysGlyGlyLysGlyValLeuLeuAspProSerTrpValIleThrAlaAla 64

Qy   223 CACTGCAAGATGAATGAGTACACCGTGCACCTGGGAGTGATACCGTG-----GGGCAC 276
             ||||||   |||   |||   |||   :: |||||||||
Db    65 HisCysSerSerAsnAsnTyrGlnValTrpLeuGlyArgAsnAsnLeuGluAspGlu 84

Qy   277 AGGAGAGTCAGAGGATCAAGGCCTCGAAGTCATTCCGCACCCCGGCTACTCCACA--- 333
             ||||:|||   |||||||   |||   :::||
Db    85 ProPheAlaGlnHisArgLeuValSerGlnSerPheProHisProAspTyrLysProPhe 104

Qy   334 ------------------------CAGACCCCATGTTAATGACCTCATGCCGTG 363
                                     ||||||||||||||||||||:::
Db   105 LeuMetArgAsnHisThrArgLysProGlyAspAspHisSerAsnAspLeuMetLeuLeu 124

Qy   364 AAGCTCAATAGCCAGGCCAGTCGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGC 423
             |||   |||||   :::||||  :::|||||||
Db   125 HisLeuSerGlnProAlaAspIleThrArgLysValIleValIleAspLeuProThrGlu 144

Qy   424 TGCGAACCCCTGGAACCACCTGTCTCCGGCTGGGCACTACCACGAGCCAGAT 483
             |||||||||:::||||||||||||||||||||||:::||
Db   145 GluProLysValGlySerThrCysLeuAlaSerGlyTrpGlySerThrLysProLeuIle 164

Qy   484 GTGACCTTTCCCTGCATGGTGGATGTCAAGTCATCATCTCCCCCAGGAGACTGC 543
             |||||||||||||||||||||||   |||||   :::||
Db   165 TrpGluPheProAspAspLeuGlnCysValAsnIleHisLeuLeuSerAsnGluLysCys 184

Qy   544 ACGAAGGTTACAAGGACTTACTGGAAAATTCCATGCTGTGGCGTCATCCCGACTCC 603
             |||||||||||:::   :::||| |||||||   :::
Db   185 IleLysAlaTyrLysGluLysValThrAspLeuMetLeuCysAlaGlyGluLeuGluGly 204

Qy   604 AAGAAAAACGCCTGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAA 663
             :::   ||||||||||||||||||||||||||||||||||||:::
Db   205 GlyLysAspThrCysThrGlyAspSerGlyGlyProLeuLeuCysAspGlyValLeuGln 224

Qy   664 GGTCTGGTCTCTGGGACCTTCCCTGCGGCGGAACTTCCCTTGCGGCCAACCCAATGACCAGGAGTCTACACT 723
             |||||:::  |||||||||||   |||||||||||||||||||||||||:::
Db   225 GlyIleThrSerTrpGlySerValProCysAlaLysThrAsnMetProAlaIleThrTyrThr 244

Qy   724 CAAGTGTGCAAGTTCACCAAGTGGATAAATGACACCATGAAAAAGCAT 771
             :::::  |||||||:::   ||||||||||:::::||
Db   245 LysLeuIleLysPheThrSerTrpIleLysGluValMetLysGluAsn 260

RESULT 10
A25606
tissue kallikrein (EC 3.4.21.35) submandibular precursor - mouse
N;Alternate names: glandular kallikrein; kininogenin
C;Species: Mus musculus (house mouse)
C;Date: 19-Nov-1988 #sequence_revision 30-Jun-1991 #text_change 09-Jul-2004
C;Accession: A25606; S06661
R;van Leeuwen, B.H.; Evans, B.A.; Tregear, G.W.; Richards, R.I.
J. Biol. Chem. 261, 5529-5535, 1986
A;Title: Mouse glandular kallikrein genes. Identification, structure, and expression of
A;Reference number: A25606; MUID:86168299; PMID:3007510
A;Accession: A25606
A;Molecule type: DNA
A;Residues: 1-261 <VAN>
A;Experimental source: renal
A;Note: the authors translated the codon GTC for residue 57 as Ala, CTC for residue 105
R;Murakami, K.; Ikigai, H.; Nagumo, N.; Tonita, M.; Shimamura, T.
FEBS Lett. 257, 400-402, 1989
A;Title: A cytocidal tissue kallikrein isolated from mouse submandibular glands.
A;Reference number: S06660; MUID:90060365; PMID:2583286
A;Accession: S06661
```

A;Molecule type: protein
A;Residues: 165-174 <MDR>
C;Genetics:
A;Gene: mGK-6
A;Introns: 16/1; 69/2; 165/2; 165/1, 210/3
C;Superfamily: trypsin; trypsin homology
C;Superfamily: hydrolase; serine proteinase
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-261/Product: tissue kallikrein #status predicted <MAT>
F;25-253/Domain: trypsin homology <TRY>
F;65,120,213/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.:   2.74e-39   Length: 261
Score: 551.00   Matches: 105
Percent Similarity: 58.46%   Conservative: 47
Best Local Similarity: 40.38%   Mismatches: 92
Query Match: 30.96%   Indels: 16
DB: Gaps: 3

US-09-905-083A-30 (1-969) x A25606 (1-261)

Qy   40 CTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAGAAGAGCCCAG------ 93
            ::||| ||||||||||| |||||||| :: ||||
Db   1 MetArgPheLeuIleLeuPheLeuAlaLeuSerLeuGlyGlyIleLeuAlaAspAlaProPro 20

Qy   94 --GGTGACAAGATTATTGATGGGCGCCCATGTGCAAGAGAGGCTCCCACCCATGCCAGGTG 150
            ::|||::: ||| || :||| ||||| ||||||
Db   21 ValGlnSerArgIleValIleGlyPheAsnCysGluLysSerGlnProTrpGlnVal 40

Qy   151 GCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGGCGGTCCTGGTCAATGAGCGGCTGGGTG 210
            ||||||||||| |:: |||:||||::|||:|||||
Db   41 AlaValTyrArgPheThrLysTyrGlnCysGlyGlyIleLeuLeuAsnValAsnTrpVal 60

Qy   211 CTCACTGCCGCCCACTGCAAGATGAGTACACCGTGCACCTGGGCAGTGATACGCTG 270
            ||||||||||||| ::||||:|||||||||||||||||
Db   61 LeuThrAlaAlaHisCysHisAsnAspLysTyrGlnValThrLeuGlyLysAsnAsnPhe 80

Qy   271 ------GGGCACAGGAGAGCTCAGAGGATCAAGGCCTCAAGTCATTCCGCCACCCCGGC 324
               ::: ||||| ||| |||||| ||||
Db   81 LeuGluAspSerAlaGlnHisHisArgLeuValSerLysAlaIleProHisProAsp 100

Qy   325 TACTCC-------------------ACACAGACCCATGTTAATGAC 351
            ::::| || |||||
Db   101 PheAsnMetSerLeuLeuAsnGluHisThrProGlnProGluAspAspTyrSerAsnAsp 120

Qy   352 CTCATGCTCGTGAAGCTCAATAGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGG 411
            ||||||| |||:||||||:||| ::::|| |||||:||
Db   121 LeuMetLeuArgLeuLysLysProAlaAspIleThrAspValValLysProIleAsp 140

Qy   412 CTGCCCTCCCGCTGCGACACCCCCTGCGACACCCTGTACTGTCTCCGGCTGGGGCCACTACC 471
            |||||||| ||||| |:::||||||
Db   141 LeuProThrGluProLysGluSerGlySerThrCysLeuAlaSerGlyTrpGlySerIle 160

Qy   472 ACGAGCCCAGATGTGACCTTCCCTCGACCTGGCGCTGGATGTCAAGTCCATCTCC 531
            ||| ::: |||||| ||| :::||||
Db   161 ThrProValLysTyrGluTyrProAspAspLeuGlnCysValAsnLeuLysLeuLeuPro 180

Qy   532 CCCCAGGACTGCACGAAGGTTTACAAGGACTTACTGGGAAAATTCCATGTGCGCTGGC 591
            ::::||||||||:::   ::   :::|||
Db   181 AsnGluAspCysAlaLysAlaAlaHisIleGluLysValThrAspAspMetLeuCysAlaGly 200

Qy   592 ATCCCGACTCCAAGAAAAACGGCTGCAATGGTGACCTCAGGGGACCGTTGGTGTGCAGA 651
            ::: :|||||| ||| |||||:||::
Db   201 AspMetAspGlyGlyLysAspThrCysAlaGlyAspSerGlyGlyProLeuIleCysAsp 220

Qy   652 GGTACCCTGCAAGGTCTGTGTCCTGCTGGGGAACTTTCCCTGCCGGCCAACCCAATGACCCA 711
            ||| ::::||||| ||||:||
Db   221 GlyValLeuGlnGlyIleThrSerTrpGlyProGlyCysGlyLysProAsnValPro 240

Qy   712 GGAGTCTACACTGTCAAGTGTCCAGAGGATCAAGCTGGATAAATGACACCATGAAAAAGCAT 771
            |||||:::||| ::||| ::::||||||
Db   241 GlyIleTyrThrArgValLeuAsnPheAsnTrpIleArgThrMetAlaGluAsn 260

RESULT 11
A37938
tissue kallikrein (EC 3.4.21.35), prostatic, precursor - dog
N;Alternate names: kallikrein homolog
C;Species: Canis lupus familiaris (dog)
C;Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 09-Jul-2004
C;Accession: A37938; S00613; A61555
R;Chapdelaine, P.; Gauthier, E.; Ho-Kim, M.A.; Bissonnette, L.; Tremblay, R.R.; Dube, J.
DNA Cell Biol. 10, 49-59, 1991
A;Title: Characterization and expression of the prostatic arginine esterase gene, a cani
A;Reference number: A37938; MUID:9119675; PMID:1991049
A;Accession: A37938
A;Molecule type: DNA
A;Residues: 1-260 <CHA>
A;Cross-references: UNIPROT:P09582; GB:M63669; NID:g163906; PIDN:AAA30831.1; PID:g163907
R;Chapdelaine, P.; Potvin, C.; Ho-Kim, M.A.; Tremblay, R.R.; Dube, J.Y.
Mol. Cell. Endocrinol. 56, 63-70, 1988
A;Title: Androgen regulation of canine prostatic arginine esterase mRNA using cloned cDN
A;Reference number: A61555; MUID:88225749; PMID:3371547
A;Accession: A61555
A;Molecule type: mRNA
A;Residues: 105-260 <CH3>
A;Note: authors translated the codon GAG for residue 67 as Gln, and TTG for residue 97 a
C;Comment: This glandular kallikrein is not known to have a physiological substrate but
C;Superfamily: trypsin; trypsin homology
C;keywords: extracellular protein; glycoprotein; hydrolase; serine proteinase
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-24/Domain: propeptide #status predicted <PRO>
F;25-252/Domain: trypsin homology <TRY>
F;25-97/Domain: tissue kallikrein, prostatic, light chain #status predicted <LCH>
F;25-97,98-260/Product: tissue kallikrein, prostatic #status predicted <MAT>
F;98-260/Domain: tissue kallikrein, prostatic, heavy chain #status predicted <HCH>
F;65,119,212/Active site: His, Asp, Ser #status predicted
F;79/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.:   5.49e-39   Length: 260
Score: 547.50   Matches: 108
Percent Similarity: 56.75%   Conservative: 35
Best Local Similarity: 42.86%   Mismatches: 96
Query Match: 30.76%   Indels: 13
DB: Gaps: 3

US-09-905-083A-30 (1-969) x A37938 (1-260)

Qy   49 TTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAAGAGCCCAG---GGTGACAAGATT 105
            |||||||:::||||||| ||||
Db   6 LeuCysLeuAlaMetSerLeuGlyTrpThrGlyAlaGluProHisPheGlnProArgIle 25

Qy   106 ATTGATGGGCGCCCATGTGCAAGAGAGGCTCCCACCCATGCCAGGTGCCCTGCTCAGTGGC 165
            |||:||||||||||||||::::||||||
Db   26 IleGlyGlyArgGluCysGluLysAsnSerGlnProTrpGlnValAlaValTyrHisAsn 45

Qy   166 AATCAGCTCCACTGCGGAGGCGGTCCTGGTCAATGAGCGGCTCACTGCCGCCCAC 225
            ::: |||||||||:|||||||
Db   46 GlyGluPheAlaCysGlyGlyValLeuValHisProGluTrpValLeuThrAlaAlaHis 65

Qy   226 TGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACGCTGGGACAGG------ 279
               |||::: ||||||||||||||
Db   66 CysAlaAsnSerAsnCysGluValTrpLeuGlyArgHisAsnLeuSerGluGluAsp 85

Qy   280 AGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGCCACCCCGGCTACTCCA------ 333
            |||:::|||| ::|:||
Db   86 GluGlyGlnLeuValGlnValArgLysSerPheIleHisProLeuTyrLysThrLysVal 105

```
Query Match:   30.42%        Indels:  11
DB:            1             Gaps:    5

US-09-905-083a-30 (1-969) x TRRT2 (1-246)

Qy   52 CTGCTATCCTTAGCCTTG-------GAAACTCAGCAGGAGAGAAGCCCAGGTGACAAG 102
        |||||  ||||||||          ||   :||||||||||||||||||||||
Db    4 LeuLeuPheLeuAlaLeuValGlyAlaAlaAlaAlaPheProValAspAspAspLys  23

Qy  103 ATTATTGATGGCGCCCACAGTGTCCAACCATGGCAAGAGGCTTCCCACCCATGGCCCTGCTCAGT 162
        |||        |||              :::      |||||||||||||||
Db   24 IleValGlyGlyTyrThrCysGlnGluAsnSerValProTyrGlnValSerLeuAsnSer  43

Qy  163 GGCAATCAGCTCCACTGCGGAGGGCTCCTGTCAATGAGCGGCTGGGTGCTCACTGCGCC 222
        |||                |||:::
Db   44 GlyTyrHisPhe---CysGlyGlySerLeuIleAsnAspGlnTrpValValSerAlaAla  62

Qy  223 CACTGCAAGATGAATGAGTACACCGTGCACTGCCGACCTGGCAGTGATACGCTG----- 270
        |||     ::::       |||     |||
Db   63 HisCysTyrLysSerArgIleGlnValArgLeuGlyGluHisAsnIleAsnValLeuGlu  82

Qy  271 GGCGACAGGAGAGTCAGAGGATCAAGGCCTCGAGTCATTCCGCACCCCGGCTACTCC 330
        |||        |||             |||   :::
Db   83 GlyAspGlu------GlnPheIleAsnAlaAlaLysIleLeuLysHisProAsnPheAsp 100

Qy  331 ACACAGACCCATGTTAATGACCTCGTCGAAGCTCAATAGCCAGGCCAGGCTGTCA 390
        :::||             |||||||||||||
Db  101 ArgLysThrLeuAsnAsnAspIleMetLeuIleLysLeuSerSerProValLysLeuAsn 120

Qy  391 TCCATGGTGAAGAAAGTCAGGCTGCCTCCCGCTGCGAACCCCTGGAACCACCTGTACT 450
        :::||||            ||     |||:::
Db  121 AlaArgValAlaThrValAlaLeuProSerSerCysAlaProAlaGlyThrGlnCysLeu 140

Qy  451 GTCTCCGGCTGGGGACTACCACGAGCCCAGATGACCTTTCCCTGACCTGTCATGTGC 510
        :::|||            |||           ||      |||:::
Db  141 IleSerGlyTrpGlyAsnThrLeuSerSerGlyValAsnGluProAspLeuLeuGlnCys 160

Qy  511 GTGGATGTCAAGCTCCATCTCCCCCAGGACTTACAAGGACTTACGGAA 570
        :::||            |||||
Db  161 LeuAspAlaProLeuLeuProGlnAlaAspCysGluAlaSerTyrProGlyLysIleThr 180

Qy  571 AATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAACGCCTGCAATGGTGACTCA 630
        :::||:::||       |||              |||||:::
Db  181 AspAsnMetValCysValGlyPheLeuGluLeuGlyLysGlyLysAspSerCysGlnGlyAspSer 200

Qy  631 GGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGTGTCCTGGGAACTTTCCCT 690
        |||||          |||||                 |||:::
Db  201 GlyGlyProValValCysAsnGlyGluLeuGlnGlyIleValSerTrpGly---TyrGly 219

Qy  691 TGCGGCCACCAATGACCACCAGGAGTCTCACACCAAGAGTGCAAGTGCAAGTGGATA 750
        |||          :::||::::::           |||     :::||||
Db  220 CysAlaLeuProAspAsnProGlyValTyrThrLysValCysAsnTyrValAspTrpIle 239

Qy  751 AATGACACCATG 762
        ||||:::
Db  240 GlnAspThrIle 243
```

```
Qy  334 -----------CAGACCCATGTTAATGACCTCGTCGAAGCTC 369
                :::                ::   |||||||||||:::
Db  106 ProArgAlaValIleArgProGlyGluArgGlyAspArgSerHisAspLeuMetLeuHisLeu 125

Qy  370 AATAGCCAGGCCAGGCTGTCATCCATGTGGAAGAAAGTCAGGAGAAAGTCGGCCTGCTGCCCTCCCGCTGCGAA 429
        ||||:::::||:::::     |||:::                  |||:::
Db  126 GluGluProAlaLysIleThrArgLysLysAlaValArgValMetAspAspLeuMetAspProLysGluPro 145

Qy  430 CCCCTGGAACCACTGTACTGGGCTGGGCACTGGGAACTCCACCAGAGACCCAGATGTGACC 489
        ||| |::::||          ||||||:::
Db  146 ProLeuGlySerThrCysTyrValSerGlyTrpGlyTyrValSerGlyTyrThrAspProGluThrArgProThrAspProGluThrArgArgThrLeuAspProArgTyrIlePhe 165

Qy  490 TTTCCCTGACCTGTCGCTGGATGTGCAAGGTGTCAGCTCATCTCCCCCAGGACTTGCACGCAAG 549
        ||:::              |||||:::||
Db  166 HisProGlySerLeuGlnGlnCysValAspLeuLeuLeuSerLeuLeuSerAlaAsnGlnCysAlaAlaLys 185

Qy  550 GTTTACAAGGACTTACTGGAAAAATTCCATGCTGTGCGCTGGCCAATCCCCGACTCCCAAGAAA 609
        |||||:::         ||||||      |||      ::   :::  |||||
Db  186 ValTyrThrGlnLysValThrLysPheMetLeuCysAlaGlyGlyLeuGlyLysGlyGlyLysLysLys 205

Qy  610 AACCCCTGCAATGGTGACTTCAGGGGGACCGTTGGTGTGCGCCAACCCAATGACCAGAGTCAAGGTCTG 669
        :::          |||||||||:::::           |||::                :::|||||
Db  206 AspThrCysLysGlyAspSerGlyGlyProLeuIleCysAspGlyGluLeuValGlyIleGlyGlyIleValGly 225

Qy  670 GTGTCCTGGGAACTTTCCCTGCGGCCAACCAATGACCAGGAGTCTCACACTCAAGTG 729
        |||||                |||||:::::           :::||:::|||||
Db  226 ThrSerTrpGlyAlaThrProCysGlyGlnLysProSerLeuTyrThrArgValThrArgVal 245

Qy  730 TGCAAGTTCACCAAGTGGATAAAATGACACCATGAAA 765
        |||||||              ||||||||||||
Db  246 MetProHisLeuMetTrpIleLysTyrAspThrMetLys 257
```

RESULT 12

TRRT2

trypsin (EC 3.4.21.4) II precursor - rat

Alignment Scores:

Pred. No.:		Length:	246
Score:	1.79e-38	Matches:	108
	541.50	Conservative:	38
Percent Similarity:	59.84%	Mismatches:	87
Best Local Similarity:	44.26%		

RESULT 13

S55067

trypsin (EC 3.4.21.4) I precursor, pancreatic - chicken

```
A;Accession: S72345
A;Molecule type: DNA
A;Residues: 1-248 <WAN2>
A;Experimental source: clone 1-P38
A;Accession: S55065
A;Molecule type: mRNA
A;Residues: 1-9,'V',11-12,'T',14-102,'A',104-214,'I',216-248 <WAN3>
A;Cross-references: EMBL:U15155; NID:g603902; PIDN:AAA79912.1; PID:g603903
A;Experimental source: clone P1
A;Accession: S72346
A;Molecule type: DNA
A;Residues: 1-9,'V',11-12,'T',14-102,'A',104-214,'I',216-248 <WAN4>
A;Cross-references: GB:U15155; NID:g603902; PIDN:AAA79912.1; PID:g603903
A;Experimental source: clone P1
R;Wang, K.
submitted to the EMBL Data Library, September 1994
A;Reference number: S71155
A;Accession: S71155
A;Molecule type: mRNA
A;Residues: 1-102,'A',104-248 <WAN5>
A;Experimental source: clone 1-P38
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-25/Domain: activation peptide #status predicted <APT>
F;26-248/Product: trypsin I #status predicted <MAT>
F;26-241/Domain: trypsin homology <TRY>
F;65,109,202/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.:
Score:                 5.33e-38          Length:          248
                       536.00            Matches:         104
Percent Similarity:    58.61%            Conservative:    39
Best Local Similarity: 42.62%            Mismatches:      97
Query Match:           30.11%            Indels:          4
DB:                                      Gaps:            2

US-09-905-083A-30 (1-969) x S55067 (1-248)

Qy   46 ATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAAGAAGCCCAGGTGACAAGATT 105
              :::      :   ::||                 :||      ||||||
Db    7 ValAlaPheLeuGlyValAlaValAlaAlaPheProIleSerAspGluAspProLysIle 26

Qy  106 ATTGATGGGCGCCCCATGTGCAAGAGGCTTCCACCCATGCAGGTGGCCCTGCTTCAGTGGC 165
                 ||||||         :::       ||||||  ::|||:|||   |||
Db   27 ValGlyGlyTyrSerCysAlaArgSerAlaAlaProTyrGlnValSerLeuAsnSerGly 46

Qy  166 AATCAGCTCCACTGCGGAGGCGGTCCTGGTCAATGAGCCGGTGCGTCCACTGCCGCCCAC 225
               :::|||||||   ||||||||||||||:::|||:|||:||::|||   ||
Db   47 TyrHisPhe--CysGlyGlySerLeuIleSerSerGlnTrpValLeuSerAlaAlaHis 65

Qy  226 TGCAAGATGAATGAATGAGTAGTACACCGTGCACCTGGGCAGTGATTACGGTGGGC-----GACAGG 279
               :::         :::|||||||   ||||||||:::       |||||
Db   66 CysTyrLysSerSerIleGlnIleValLysLeuGlyGluTyrAsnLeuAlaAlaAspGly 85

Qy  280 AGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGCCACCCCGGCTACTCCACCAGACC 339
               :::|||||||         :::|||::       |||||:::|||
Db   86 SerGluGlnThrIleSerSerLysValIleArgHisSerGlyTyrAsnSerAsnThr 105

Qy  340 CATGTTAATGACCTATGCTCGTGAAGCTCAATAGCCAGGCCAGGCTGTCATCCATGGTG 399
               ||||||:::||||||:::|||||||:::       ||||||
Db  106 LeuAsnAsnAspIleMetLeuIleLysLeuSerlysAlaAlaThrLeuAsnSerTyrVal 125

Qy  400 AAGAAAGTCAGGCTGCCCTCCCGCTGCGAACCCCTGGAACCACCTGTACTGTCTCCGGC 459
               |||::       :::|||||||   |||||||||
Db  126 AsnThrValProLeuProThrSerCysValThrAlaGlyThrThrCysLeuIleSerGly 145

Qy  460 TGGGGCACTACCACGAGCCCAGATGTGACCTTTCCCTCAGCCTCCATGTGCGTGGATGTC 519
               |||||:::|||::       :::|||||||   :::|||
Db  146 TrpGlyAsnThrLeuSerSerGlySerLeuTyrProAspValLeuGlnCysLeuAsnAla 165

Qy  520 AAGCTCATCTCCCCCCAGGACTGCACGAGGGTTTACAAGGGACTTACTGGAAAATTCCATG 579
               |||:::|||       :::|||:||
Db  166 ProValLeuSerSerGlnCysSerSerAlaTyrProGlyArgIleThrSerAsnMet 185

Qy  580 CTGTGCGCTGGCATCCCGACTCCAAGAAAAACGCCTGCAATGTGACTCAGGGGACCG 639
               :::|||      |||      :::||||||||||||:::::::|||
Db  186 IleCysIleGlyTyrLeuAsnGlyLysLysAspSerCysGlnGlyAspSerGlyGlyPro 205

Qy  640 TTGGTGTGCAGAGGTACCCTGCCAAGGTCTGGTGTCCTGGGAACTTCCCTTGCCGCCAA 699
               ||||||||||:::|||||||:::|||||||||||||:::
Db  206 ValValCysAsnGlyGlnLeuGlnGlyPheValSerTrpGly--IleGlyCysAlaGln 224

Qy  700 CCCAATGACCCAGGAGTCTACACTCAAGTGTGCAAGTTCACCAAGTGGATAAATGACACC 759
               :::      :||||||:::|||||||||:::|||
Db  225 LysGlyTyrProGlyValTyrThrLysValCysAsnTyrValSerTrpIleLysThrThr 244

Qy  760 ATGAAAAAGCAT 771
               ||| :::
Db  245 MetSerSerAsn 248

RESULT 14
TRBOTR
trypsin (EC 3.4.21.4) precursor - bovine
N;Contains: trypsinogen
C;Species: Bos primigenius taurus (cattle)
C;Date: 24-Apr-1984 #sequence_revision 28-Feb-1986 #text_change 18-Jul-1997
C;Accession: A90164; A00946; S08774
R;Mikes, O.; Holeysovsky, V.; Tomasek, V.; Sorm, F.
Biochem. Biophys. Res. Commun. 24, 346-352, 1966
A;Title: Covalent structure of bovine trypsinogen. The position of the remaining amides.
A;Reference number: A90164; MUID:67168848; PMID:5967094
A;Accession: A90164
A;Molecule type: protein
A;Residues: 1-57,'Q',59-67,'Q',69-150,'N',152-176,'N',178-229 <MIK>
R;Hartley, B.S.
Philos. Trans. R. Soc. Lond. B257, 77-87, 1970
A;Reference number: A93755
A;Contents: annotation; revisions
R;Titani, K.; Ericsson, L.H.; Neurath, H.; Walsh, K.A.
Biochemistry 14, 1358-1366, 1975
A;Title: Amino acid sequence of dogfish trypsin.
A;Reference number: A00950; MUID:75146445; PMID:1092332
A;Contents: annotation; revisions
A;Note: the sequence agrees with that shown
J. Mol. Biol. 98, 693-717, 1975
A;Title: The refined crystal structure of bovine beta-trypsin at 1.8 angstrom resolution
A;Reference number: A92954; MUID:76072097; PMID:512
A;Contents: annotation; X-ray crystallography; binding sites for calcium, substrate, and
C;Comment: Trypsinogen is synthesized in the acinar cells of the pancreas.
C;Comment: Autocatalytic cleavage after Lys-6 leads to beta-trypsin by releasing a termi
s pseudotrypsin. A cleavage may also occur after Arg-105.
C;Superfamily: hydrolase; pancreas; protein digestion; serine proteinase; zymogen
F;1-229/Product: trypsinogen #status experimental <ZYM>
F;1-6/Domain: activation peptide #status experimental <APT>
F;7-222/Domain: trypsin homology <TRY>
F;7-131,132-229/Product: alpha-trypsin #status experimental <MPT>
F;6-7/Cleavage site: Lys-Ile (enteropeptidase) #status experimental
F;13-143,31-47,115-216,122-189,154-168,179-203/Disulfide bonds: #status experimental
F;46,90,183/Active site: His, Asp, Ser #status experimental
F;58,60,63,68/Binding site: calcium (Glu, Asn, Val, Glu) #status experimental
F;131-132/Cleavage site: Lys-Ser (autolytic) #status experimental

Alignment Scores:
Pred. No.:
Score:                 6.47e-38          Length:          229
                       535.00            Matches:         100
Percent Similarity:    60.71%            Conservative:    36
Best Local Similarity: 44.64%            Mismatches:      84
Query Match:           30.06%            Indels:          4
DB:                                      Gaps:            3

US-09-905-083A-30 (1-969) x TRBOTR (1-229)
```

Biochim. Biophys. Acta 866, 1-14, 1986
A;Title: Immunological identification of rat tissue kallikrein cDNA and characterization
A;Reference number: A25137; MUID:86131678; PMID:3004582
A;Accession: A25137
A;Molecule type: mRNA
A;Residues: 115-265 <GER>
R;Inoue, H.; Fukui, K.; Miyake, Y.
J. Biochem. 105, 834-840, 1989
A;Title: Identification and structure of the rat true tissue kallikrein gene expressed i
A;Reference number: JX0073; MUID:89327211; PMID:2753879
A;Accession: JX0073
A;Molecule type: DNA
A;Residues: 1-265 <INO>
A;Cross-references: GB:D00448; NID:g220792; PIDN:BAA00346.1; PID:g220794
A;Experimental source: kidney
R;Ashley, P.L.; MacDonald, R.J.
Biochemistry 24, 4512-4520, 1985
A;Title: Kallikrein-related mRNAs of the rat submaxillary gland: nucleotide sequences of
A;Reference number: A23863; MUID:86051477; PMID:2998455
A;Accession: A23863
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-265 <ASH>
A;Cross-references: GB:M11563; NID:g205029; PIDN:AAA41464.1; PID:g205030
A;Experimental source: submaxillary gland
R;Wines, D.R.; Brady, J.M.; Pritchett, D.B.; Roberts, J.L.; MacDonald, R.J.
J. Biol. Chem. 264, 7653-7662, 1989
A;Title: Organization and expression of the rat kallikrein gene family.
A;Reference number: A33359; MUID:89214217; PMID:2708383
A;Accession: A33359
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 5-265 <WIN>
A;Cross-references: GB:M23874; GB:J04701; GB:M23875; GB:M23876; NID:g205007; PID:g205007; PIDN:AAA414
C;Comment: The kallikreins liberate lysyl-bradykinin, a vasoactive decapeptide, from kin
C;Comment: The protein presumably assumes the two-chain form by cleavage between residu
C;Comment: Tissue kallikreins cleave Met-Lys and Arg-Ser bonds in kininogen to release 1
C;Genetics:
A;Introns: 20/1; 73/2; 169/1; 214/3
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; pancreas; serine proteinase; zymogen
F;1-17/Domain: signal sequence #status predicted <SIG>
F;18-28/Domain: activation peptide #status predicted <APT>
F;29-265/Product: tissue kallikrein, pancreatic #status predicted <MPT>
F;23-257/Domain: trypsin homology <TRY>
F;35-177,54-70,156-223,188-202,213-238/Disulfide bonds: #status predicted
F;69,124,217/Active site: His, Asp, Ser #status predicted

Alignment Scores:
Pred. No.:   Length: 265
Score: 6.53e-38   Matches: 104
Percent Similarity: 535.00   Conservative: 44
Best Local Similarity: 57.59%   Mismatches: 93
Query Match: 40.47%   Indels: 16
DB: 30.06%   Gaps: 3

US-09-905-083a-30 (1-969) x KQRTP (1-265)

Qy 49 TTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGGAGGAGGAGAGCCCAG--------GGTGAC 99
        ||||:::||| |||||||| :::|||
Db   8 LeuLeuLeuPheLeuAlaLeuSerLeuGlyLysArgAsnAspAlaAlaProProValGlnSer 27

Qy 100 AAGATTATTGATGGGCCCCATGTGCAAGAGGCTCCACCATGGCAGGTGGCCCTGCTC 159
        :::::::||| ||| ||||||||||| |:::
Db 28 ArgValValGlyValGlyTyrAsnCysGluMetAsnSerGlnProTrpGlnValAlaValTyr 47

Qy 160 AGTGGCAATCAGCTCCACTGCGGAGGGCGTCCTGTCAATGAGCCTGGGTGCTCACTGCC 219
        |||||| ||| |||||
Db 48 TyrPheGlyGluTyrLeuCysGlyGlyValLeuIleAspProSerTrpValIleThrAla 67

Qy 220 GCCCACTGCAAGATGCAAGATGAATGAAGTACACCGTGCACCTGGGCAGTGATACGCTGGGGCACAGG 279
        |||||||||| :::   ::| ||||||
Db 68 AlaHisCysAlaThrAspAsnTyrGlnValTrpLeuGlyArgAsnAsnLeuTyrGluAsp 87

Qy 97 GACAAGATTATTGATGGGCCCCATGTGCAAGAGGCTCCACCATGGCAGGTGGCCCTG 156
        ||||||||||::: ||| ::: |||:||||||
Db   5 AspLysIleValGlyValGlyTyrThrCysGlyAlaAsnThrValProTyrGlnValSerLeu 24

Qy 157 CTCAGTGGCAATCAGCTCCACTGCGGAGGGCGTCCTGGGTCTCACT 216
        |||||| ||:::||||::::::
Db 25 AsnSerGlyTyrHisPhe---CysGlyGlySerLeuIleAsnSerGlnTrpValValSer 43

Qy 217 GCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGACGGCTGATACGCTGGGC--- 273
        ||| ::: ||| |||
Db 44 AlaAlaHisCysTyrLysSerGlyIleValArgLeuGlyGluGluAsnIleAsnVal 63

Qy 274 ---GACAGGAGAGTCAGAGGATCAAGGCCTGAAGTCATTCCGCCACCCGGCTACTCC 330
         ::: ||||||
Db 64 ValGluGlyLysAsnGluGlnPheIleSerAlaSerLysLysSerIleValHisProSerTyrAsn 83

Qy 331 ACACAGACCCCATGTTAATGACCTCGTGAGAGCTCAATAGCCAGGCCAGGCTGTCA 390
        ::: |||| |||||||:::||| ||
Db 84 SerAsnThrLeuAsnAsnAspIleMetLeuIleLysLeuLysSerAlaAlaSerLeuAsn 103

Qy 391 TCCATGGTGAAGAAAGTCAGGCTGCCCTCCCCGCTGCGAACCCCCTGGAACCACCTGTACT 450
        ||| ::: ||| |||
Db 104 SerArgValAlaSerIleSerLeuProThrSerCysAlaSerAlaGlyThrGlnCysLeu 123

Qy 451 GTCTTCCGGCTGGGGCACTACCACGAGCCCAGATGACCTTTCCCTCTGACCTCATGTGC 510
        :::||||| ||| ||:::||
Db 124 IleSerGlyTrpGlyAsnThrLysSerSerGlyThrSerTyrProAspValLeuLeuCys 143

Qy 511 GTGGATGTCAAGCTCATCTCCCCCAGACTGCACGAGGTTTACAAGGACTTACTGGAA 570
         ::::|||||
Db 144 LeuLysAlaProIleLeuSerAspSerSerCysLysSerSerAlaTyrProGlyInIleThr 163

Qy 571 AATTCCATGCTGTGCGGCTGGCATCCCGACTCCAAGAAAAACGCCTGCAATGTGACTCA 630
        :::::||| ||||||| ||:::||| ||||||
Db 164 SerAsnMetPheCysAlaGlyTyrLeuGluGlyGlyLysAspSerCysGlnGlyAspSer 183

Qy 631 GGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCTGGGGAACTTTCCCT 690
        |||||||| |||||| ||:::||
Db 184 GlyGlyProValCysSerGlyLysLeuGlnGlyIleValSerTrpGlySer---Gly 202

Qy 691 TGCCGGCCAACCCAATGACCCAGGAGTCTACACTCCAAGTGTGCAAGTTCACCAAGTGGATA 750
        ||| |||||||||||||||::::||||
Db 203 CysAlaGlnLysAsnLysProGlyValTyrThrLysValCysAsnTyrValSerTrpIle 222

Qy 751 AATGACACCATG 762
        ||:::
Db 223 LysGlnThrIle 226

RESULT 15
KQRTP
tissue kallikrein (EC 3.4.21.35) precursor - rat
N;Alternate names: glandular kallikrein; kininogenin; true tissue kallikrein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 05-Apr-1983 #sequence_revision 05-Apr-1983 #text_change 09-Jul-2004
C;Accession: A00944; A41429; A25137; JX0073; A23863; A33359
R;Swift, G.H.; Dagorn, J.C.; Ashley, P.L.; Cummings, S.W.; MacDonald, R.J.
Proc. Natl. Acad. Sci. U.S.A. 79, 7263-7267, 1982
A;Title: Rat pancreatic kallikrein mRNA: nucleotide sequence and amino acid sequence of
A;Reference number: A00944; MUID:83117659; PMID:6961406
A;Accession: A00944
A;Molecule type: mRNA
A;Residues: 1-265 <SWI>
A;Cross-references: UNIPROT:P00758
A;Experimental source: pancreatic
C;Date: 05-Apr-1983 #sequence_revision
R;Kato, H.; Nakanishi, E.; Enjyoji, K.; Hayashi, I.; Oh-ishi, S.; Iwanaga, S.
J. Biochem. 102, 1389-1404, 1987
A;Title: Characterization of serine proteinases isolated from rat submaxillary gland: wi
A;Reference number: A41429; MUID:88198057; PMID:3482210
A;Accession: A41429
A;Status: preliminary
A;Molecule type: protein
A;Residues: 29-53,'X,'55-87 <KAT>
R;Gerald, W.L.; Chao, J.; Chao, L.

```
Qy    280 AGA------GCTCAGGAGGATCAAGGCCTCGAAGTCATTCCGCCACCCCGGCTACTCCACA 333
              |||||          |||::|||||  |||||||||:::::::
Db     88 GluProPheAlaGlnHisArgLeuValSerGlnSerPheProHisProGlyPheAsnGln 107

Qy    334 ------------------------------CAGACCCATGTTAATGACCTCATGCTC 360
                                          :::  ||||||||||||||||
Db    108 AspLeuIleTrpAsnHisThrArgGlnProGlyAspAspTyrSerAsnAspLeuMetLeu 127

Qy    361 GTGAAGCTCAATAGCCAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCC 420
          :::  ||::|||  :::::: |||||    ::: |||||
Db    128 LeuHisLeuSerGlnProAlaAspIleThrAspGlyValLysValIleAspLeuProIle 147

Qy    421 CGCTGCGAACCCCTGGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCA 480
              ||:::|||||              |||||||||||| :::
Db    148 GluGluProLysValGlySerThrCysLeuAlaSerGlyTrpGlySerIleThrProAsp 167

Qy    481 GATGTGACCTTTCCCTCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGAC 540
              :::              |||||||||              |||:::|||      :::
Db    168 GlyLeuGluLeuSerAspAspLeuGlnCysValAsnIleAspLeuSerAsnGluLys 187

Qy    541 TGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCCGAC 600
          |||  :::  ::|||:::::   :::  :::   |||||||||||||||||      ||
Db    188 CysValGluAlaHisLysGluGluValThrAspLeuMetLeuCysAlaGlyMetAsp 207

Qy    601 TCCAAGAAAAACGGCTGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTG 660
              |||::: |||       |||||||||      ||||||||||:::|||      |||
Db    208 GlyGlyLysAspThrCysLysGlyAspSerGlyGlyProLeuIleCysAsnGlyValLeu 227

Qy    661 CAAGGTCTGGTGTCCTGGGGAACTTCCCTTGCCGGCCAACCCAATGACCCAGGAGTCTAC 720
          |||||||:::    |||||||| ||||||||||||:::|||  :::::      ||||||
Db    228 GlnGlyIleThrSerTrpGlyPheAsnProCysGlyGluProLysProGlyIleTyr 247

Qy    721 ACTCAAGTGTGCAAGTTCACCAAGTGGATAAATGACACCATGAAAAAAGCAT 771
          |||:::::: |||||||     |||||||  :::  ||||||||:::::::
Db    248 ThrlysLeuIleLysPheThrProTrpIleLysGluValMetLysGluAsn 264
```

Search completed: July 14, 2005, 17:33:33
Job time : 50.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 14, 2005, 17:07:08 ; Search time 203.5 Seconds
(without alignments)
4876.705 Million cell updates/sec

```
Title:           US-09-905-083A-30
Perfect score:   1780
Sequence:        1 ggattccgggctccatggc.........aagaaacacaaaccctcag 969

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters:      3224756

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries
```

```
Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09905083/runat_14072005_140352_10610/app_query.fasta_1.1159
-DB=UniProt -QFMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09905083 @CGN_1_1_244 @runat_14072005_140352_10610 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=30 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  UniProt_03:*
         1: uniprot_sprot:*
         2: uniprot_trembl:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1364	76.6	253	1	KLK7_HUMAN	P49862 homo sapien
2	1353	76.0	253	2	Q8N5N9	Q8n5n9 homo sapien
3	1028.5	57.8	249	2	Q91VE3	Q91ve3 m thymopsin
4	988	55.5	181	2	Q8NFV7	Q8nfv7 homo sapien
5	951.5	53.5	234	2	Q9R048	Q9r048 mus musculu
6	614	34.5	260	1	NRPN_MOUSE	Q61955 mus musculu
7	612.5	34.4	293	1	KLK5_HUMAN	Q9y337 homo sapien
8	611.5	34.4	260	1	NRPN_RAT	O88780 rattus norv
9	609.5	34.2	276	2	Q8CGR6	Q8cgr6 mus musculu
10	604.5	34.0	246	2	Q6P3Z0	Q6p3z0 mus musculu
11	593.5	33.3	250	2	Q8CGR5	Q8cgr5 mus musculu
12	589	33.1	293	2	Q9D140	Q9d140 mus musculu
13	586.5	32.9	242	2	Q80VS4	Q80vs4 mus musculu
14	583	32.8	251	1	KLKE_HUMAN	Q9p0g3 homo sapien
15	583	32.8	251	2	Q6B0B9	Q6b0b9 homo sapien
16	582.5	32.7	260	1	KLK8_HUMAN	O60259 homo sapien
17	582	32.7	261	1	KLK9_MOUSE	P15949 mus musculu
18	580.5	32.6	260	2	Q8IW69	Q8iw69 homo sapien
19	578	32.5	277	1	KLKD_HUMAN	Q9ukr3 homo sapien
20	576.5	32.4	239	2	Q63275	Q63275 rattus norv
21	573.5	32.2	255	2	Q7JIG6	Q7jig6 saguinus oe
22	571.5	32.1	250	1	KLKB_HUMAN	Q9ubx7 homo sapien
23	568	31.9	261	1	KLKB_MOUSE	P15946 mus musculu
24	567.5	31.7	254	1	KLK4_HUMAN	Q9y5k2 homo sapien
25	564.5	31.6	250	2	Q63ZF2	Q63zf2 rattus norv
26	563	31.6	256	1	KLKF_HUMAN	Q9hzr5 homo sapien
27	563	31.6	261	1	KLK3_MOUSE	P00756 mus musculu
28	563	31.6	261	1	KLK8_RAT	P36374 rattus norv
29	561.5	31.5	255	2	Q6ISI0	Q6isi0 homo sapien
30	561.5	31.5	276	2	Q9QYN3	Q9qyn3 m hippostas
31	561	31.5	246	1	TRY1_RAT	P00762 rattus norv
32	560	31.5	263	1	KLKR_MOUSE	Q9jm71 mus musculu
33	559.5	31.4	255	2	Q96RQ0	Q96rq0 homo sapien
34	556	31.2	275	2	Q8IXD7	Q8ixd7 homo sapien
35	555	31.2	248	1	KLKC_HUMAN	Q9ukr0 homo sapien
36	554.5	31.2	234	2	Q9CV76	Q9cv76 mus musculu
37	553.5	31.1	249	2	Q9QYN4	Q9qyn4 mus musculu
38	552.5	31.0	260	1	ESTA_CANFA	P09582 canis famil
39	552	31.0	261	1	KLK1_MOUSE	P00755 mus musculu
40	552	31.0	261	1	KLK6_MOUSE	P15947 mus musculu
41	552	31.0	261	1	KLK7_RAT	P36373 rattus norv
42	551.5	31.0	246	2	Q6IE66	Q6ie66 rattus norv
43	551	31.0	261	2	Q8C232	Q8c232 mus musculu
44	551	31.0	276	1	KLKA_HUMAN	O43240 homo sapien
45	548	30.8	263	1	KLKO_MOUSE	Q61754 mus musculu

ALIGNMENTS

```
RESULT 1
KLK7_HUMAN
ID   KLK7_HUMAN     STANDARD;      PRT;    253 AA.
AC   P49862;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Kallikrein 7 precursor (EC 3.4.21.-) (Stratum corneum chymotryptic
DE   enzyme) (hSCCE).
GN   Name=KLK7; Synonyms=PRSS6, SCCE;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 23-53.
RC   TISSUE=Skin;
RX   MEDLINE=94308225; PubMed=8034709;
RA   Hansson L., Stroemqvist M., Baeckman A., Wallbrandt P., Carlstein A.,
RA   Egelrud T.;
RT   "Cloning, expression, and characterization of stratum corneum
RT   chymotryptic enzyme. A skin-specific human serine proteinase.";
RL   J. Biol. Chem. 269:19420-19426(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Keratinocytes;
RA   Yousef G.M., Scorilas A., Diamandis E.P.;
RT   "Molecular characterization, mapping and tissue expression of the
RT   human stratum corneum chymotryptic enzyme gene.";
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;
RA   Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
RA   Moss P., Paeper B., Wang K.;
RT   "Sequencing and expression analysis of the serine protease gene
RT   cluster located in chromosome 19q13 region.";
RL   Gene 257:119-130(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
```

RA   Hansson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,
RA   Wallbrandt P., Egelrud T.;
RT   "Epidermal overexpression of stratum corneum chymotryptic enzyme in
RT   mice; a model for chronic itchy dermatitis.";
RL   Biochem. Biophys. Res. Commun. 211:586-589(1995).
RN   [5]
RP   CHARACTERIZATION.
RX   MEDLINE=95314630; PubMed=7794273;
RA   Skytt A., Stroemqvist M., Egelrud T.;
RT   "Primary substrate specificity of recombinant human stratum corneum
RT   chymotryptic enzyme.";
RL   Biochem. Biophys. Res. Commun. 211:586-589(1995).
CC   -!- FUNCTION: May catalyze the degradation of intercellular cohesive
CC   structures in the cornified layer of the skin in the continuous
CC   shedding of cells from the skin surface. Specific for amino acid
CC   residues with aromatic side chains in the P1 position. SCCE
CC   cleaves insulin B chain at 6-Leu-|-Cys-7, 16-Tyr-|-Leu-17, 25-
CC   Phe-|-Tyr-26, and 26-Tyr-|-Thr-27. Could play a role in the
CC   activation of precursors to inflammatory cytokines.
CC   -!- TISSUE SPECIFICITY: Abundantly expressed in the skin and is
CC   expressed by keratinocytes in the epidermis. Very low levels are
CC   also seen in the brain and kidney.
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
CC   subfamily.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; L33404; AAC37551.1; -.
DR   EMBL; AF166330; AAD49718.1; -.
DR   EMBL; AF243527; AAG33360.1; -.
DR   EMBL; AF332583; AAK69624.1; -.
DR   PIR; A53968; A53968.
DR   HSSP; P00760; 1EZX.
DR   MEROPS; S01.300; -.
DR   Genew; HGNC:6368; KLK7.
DR   H-InvDB; HIX0015373; -.
DR   MIM; 604438; -.
DR   GO; GO:0008236; F:serine-type peptidase activity; TAS.
DR   GO; GO:0008544; P:epidermal differentiation; TAS.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Direct protein sequencing; Glycoprotein; Hydrolase; Serine protease;
KW   Signal; Zymogen.
FT   SIGNAL   1   22
FT   PROPEP   23   29   Activation peptide.
FT   CHAIN   30   253   Kallikrein 7.
FT   ACT_SITE   70   70   Charge relay system (By similarity).
FT   ACT_SITE   112   112   Charge relay system (By similarity).
FT   ACT_SITE   205   205   Charge relay system (By similarity).
FT   DISULFID   36   71   By similarity.
FT   DISULFID   55   ?   By similarity.
FT   DISULFID   144   211   By similarity.
FT   DISULFID   176   190   By similarity.
FT   DISULFID   201   226   By similarity.
FT   CARBOHYD   246   246   N-linked (GlcNAc...) (Potential).
SQ   SEQUENCE   253 AA; 27525 MW; 2D68B6B15A76A668 CRC64;

Alignment Scores:
Pred. No.:   6.12e-110   Length: 253

Score: 1364.00   Matches: 253
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00%   Mismatches: 0
Query Match: 76.63%   Indels: 1
DB: Gaps: 0

US-09-905-083A-30 (1-969) x KLK7_HUMAN (1-253)

```
Qy     16 ATGGCAAGATCCCTTCCTGCCCTGCGAGATCTTACTGCTCATCCTTAGCCTTGGAAACT  75
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr  20

Qy     76 GCAGGAGAAGAGCCCAGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCC  135
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21 AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaArgGlySer  40

Qy    136 CACCCATGGCAGGTGGCCCTGCTCAGTGGCAATCAGTCCACTGCGGAGGCGTCCTGGTC  195
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41 HisProTrpGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal  60

Qy    196 AATGACGCGTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTG  255
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu  80

Qy    256 GGCAGTGATACGCTGGGCGACAGGAGAGCTCAAGAGGATCAAGGCCTCGAAGTCATTCCGC  315
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81 GlySerAspThrLeuGlyAspArgArgAlaGlnGluAspGlnGlyLeuGluValIleProPro  100

Qy    316 CACCCGGCTACTCCACAGACCCATGTTAATGACCTCGTGAAGCTCAATAGC  375
              |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101 HisProGlyTyrSerThrGlnThrHisValAsnAspLeuValLysLeuAsnSer  120

Qy    376 CAGGCCAGGTGCGTCATCCATGGTGGATGTGAAGAAAGTCAGGCTGCGCCTGCCGACGAACCCCT  435
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GlnAlaArgLeuSerMetValLysLysValAlaArgLeuProSerArgCysGluProPro  140

Qy    436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC  495
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141 GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro  160

Qy    496 TCTGACCTCATGTGCGTGACTGTCAAGCTCATCTCCCCCAGGACTGACGAAGGTTTAC  555
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161 SerAspLeuMetCysValAspValLysLeuIleSerProGlnAspCysThrLysValTyr  180

Qy    556 AAGGACTTACTGGAAAATTCCATGCTGTGCGGCTGCATCCCGACTCCAAGAAAAACGCC  615
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 LysAspLeuLeuGluAsnSerMetLeuCysGlyCysIleProThrProArgLysLysAsnAla  200

Qy    616 TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCC  675
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201 CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer  220

Qy    676 TGGGGAACTTTCCCTGCGGCCAACCAATGACCCAGAGTCTACACTCAAGTGTGCAAG  735
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221 TrpGlyThrPhePheProCysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLys  240

Qy    736 TTCACCAAGTGGATAAATGACACCATGAAAAAGCATCGC  774
              |||||||||||||||||||||||||||||||||||||||
Db    241 PheThrLysTrpIleAsnAspThrMetLysLysHisArg  253
```

RESULT 2
Q8N5N9
ID   Q8N5N9   PRELIMINARY;   PRT;   253 AA.
AC   Q8N5N9;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stratum corneum chymotryptic enzyme, preproprotein.
GN   Name=KLK7;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;

```
Qy    256  GGCAGTGATACGCTGGGCGACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC  315
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     81  GlySerAspThrLeuGlyAspArgArgAlaGlnArgIleLysAlaSerLysSerPheArg  100

Qy    316  CACCCCGGCTACTCCACACAGACCACCATGTTAATGACCTCGTGAAGCTCAATAGC      375
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    101  HisProGlyTyrSerThrGlnThrThrMetLeuMetThrSerValLysLeuAsnSer    120

Qy    376  CAGGCCAGCTGTCATCCATGGTGAAGAAGTCAGGCTGCCTCCCGCTGCGAACCCCT      435
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121  GlnAlaSerCysHisProTrpTrpLysLysSerGlyCysLeuProLeuArgThrPro    140

Qy    436  GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC  495
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    141  GlyThrThrCysThrValSerGlyTrpGlyThrThrThrSerProAspValThrPhePro  160

Qy    496  TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTAC   555
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    161  SerAspLeuMetCysValAspValLysLeuIleSerProProThrAlaArgArgPheThr 180

Qy    556  AAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGACATCCCGACTCCAAGAAAAACGCC  615
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181  LysAspLeuLeuGluAsnSerMetLeuCysAlaGlyIleProAspSerLysLysAsnAla  200

Qy    616  TGCAATGGTGACTCAGGGGGACCGTGGTGTGGAGGTACCCTGCAAGGTCTGGTGTCC     675
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    201  CysAsnGlyAspSerGlyGlyProLeuValCysArgGlyThrLeuGlnGlyLeuValSer  220

Qy    676  TGGGGAACTTCCTTGCGGCCAACCCAATGACCCAGGAGTCTACACTCCAAGTGTGCAAG  735
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    221  TrpGlyThrPheProTrpGlyGlnProAsnAspProGlyValTyrThrProLysCysLys  240

Qy    736  TTCACCAAGTGGATAAATGACACCATGAAAAAGCATGC  774
                |||||||||||||||||||||||||||||||||||||
Db    241  PheThrLysTrpIleAsnAspThrMetLysLysHisArg  253
```

```
RESULT 3
Q91VE3
ID   Q91VE3              PRELIMINARY;     PRT;    249 AA.
AC   Q91VE3;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Thymopsin (kallikrein 7) (Stratum corneum chymotryptic enzyme) (Mus
DE   musculus 0 day neonate head cDNA, RIKEN full-length enriched library,
DE   clone:4833444E03 product:kallikrein 7 (chymotryptic, stratum corneum),
DE   full insert sequence) (Mus musculus 10 days neonate head cDNA, RIKEN
DE   full-length enriched library, clone:5530400G09 product:kallikrein 7
DE   (chymotryptic, stratum corneum), full insert sequence).
GN   Name=Klk7;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Thymus;
RA   Yamaguchi N.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RA   Strausberg R.;
RL   Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; BC032005; AAH32005.1; -.
DR   HSSP; P00760; 1E2X.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PRINTS; PR00089; TRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE 253 AA; 27608 MW; 2D68B6A41B22A668 CRC64;
```

```
Alignment Scores:
Pred. No.:                          5.55e-109          Length:           253
Score:                              1353.00            Matches:          252
Percent Similarity:                 99.60%             Conservative:       0
Best Local Similarity:              99.60%             Mismatches:         1
Query Match:                        76.01%             Indels:             0
DB:                                 2                  Gaps:               0

US-09-905-083A-30 (1-969) x Q8N5N9 (1-253)
```

```
Qy     16  ATGGCAAGATCCCTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACT   75
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1  MetAlaArgSerLeuLeuLeuProLeuGlnIleLeuLeuLeuSerLeuAlaLeuGluThr   20

Qy     76  GCAGGAGAAGAAGCCCAGGGTGACAAGATTATTGATGGCGCGCCCCATGTGCAAGAGGCTCC 135
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     21  AlaGlyGluGluAlaGlnGlyAspLysIleIleAspGlyAlaProCysAlaProArgGlySer 40

Qy    136  CACCCATGCCAGGTGGCCCTGCTCAGTGCAATCAGCTCCACTGCGGAGGCGGTCCTGGTC  195
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     41  HisProThrGlnValAlaLeuLeuSerGlyAsnGlnLeuHisCysGlyGlyValLeuVal   60

Qy    196  AATGAGCGGCTGGGTGCTCACTCGCCCACTGCAAGATGAATGAGTACCGTGCACCTG     255
                |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61  AsnGluArgTrpValLeuThrAlaAlaHisCysLysMetAsnGluTyrThrValHisLeu   80
```

RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=129/SvJ;
RA   Hansson L., Backman A., Ny A., Edlund M., Edholm E., Tornell J.,
RA   Wallbrandt P., Egelrud T.;
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [7]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RA   The FANTOM Consortium.
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [8]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [9]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RX   MEDLINE=20310913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake Y., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [10]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RA   Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA   Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA   Hayashida K., Hayatsu N., Hiramoto K., Hiramoto T., Hirozane T.,

RA   Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA   Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA   Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA   Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA   Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA   Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA   Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA   Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN   [11]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Head;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA   Tejima Y., Toya T., Yamamura T., Yamanaka I., Yasunishi A.,
RA   Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; AB008371; BAB55604.1; -.
DR   EMBL; BC027823; AAH27823.1; -.
DR   EMBL; AF339930; AAK69652.1; -.
DR   EMBL; AK029477; BAC26467.1; -.
DR   EMBL; AK077406; BAC36787.1; -.
DR   HSSP; P00760; 1EZX.
DR   MEROPS; S01.300; -.
DR   MGD; MGI:1346336; Klk7.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   249 AA; 27257 MW; 0D4E380F12D14F87 CRC64;

Alignment Scores:
Pred. No.:   Length: 249
Score: 9.82e-81   Matches: 185
Percent Similarity: 1028.50   Conservative: 27
Best Local Similarity: 85.48%   Mismatches: 33
Query Match: 74.60%   Indels: 3
DB: 57.78%   Gaps: 1
                       2

US-09-905-083A-30 (1-969) x Q91VE3 (1-249)

Qy   31 CTCCTGCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAAGAAGCC 90
             ||||| ||| :::|||||||||||||||||||||||
Db   5 LeuLeuSerLeuIleThrValLeuLeuSerLeuLeuAlaLeuGluThrAlaGly------ 21

Qy   91 CAGGGTGACAAGATTATTGATGCGCCCCATGTGCAAGAGGCTCCACCCATGCAGAGGTG 150
             ||||||||||::::||||||||||| || |||||||||||||||||
Db   22 GlnGlyGluArgIleIleAspGlyTyrLysCysGluGlySerHisProTrpGlnVal 41

Qy   151 GCCCTGCTCAGTGGCAATCAGCTCCACTGCGGAGCGTCCTGGTCAATGAGCGCTGGGTG 210
             |||||||||||||| |||||||||||||||||||||||||||||||
Db   42 AlaLeuLeuLysGlyAsnGlnLeuHisCysGlyGlyValLeuValAspLysTyrTrpVal 61

Qy   211 CTCACTGCGCGCCCACTGCAAGATGAATGAAGTACACCGTGCACCTGGCAGTGATACGCTG 270
             |||||||||:::||||| ||:||:| |:||||
Db   62 LeuThrAlaAlaHisCyslysMetGlyInTyrGlnIleLeuGlnLeuGlySerAspLysIle 81

Qy   271 GGGCACAGGAGAGCTCAGAGATCAAGGAATCAAGGCCTCGAAGTCATTCCGCCACCCGGCTACTCC 330
             ||||||::: |||||| |||||::||||||||||
Db   82 GlyAspGlnSerAlaGlnLysLysAlaThrLysLysSerPheArgHisProArgTyrTyrSer 101

```
Percent Similarity:      100.00%            Conservative:  0
Best Local Similarity:   100.00%            Mismatches:    0
Query Match:             55.51%             Indels:        0
DB:                                         Gaps:          0

US-09-905-083a-30 (1-969) x Q8NFV7 (1-181)

Qy   232 ATGAATGAGTACACCGTGCACCTGGGCAGTGATACGCTGGGCGACAGGAGAGTCAGAGG 291
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 MetAsnGluTyrThrValHisLeuGlySerAspThrLeuGlyAspArgArgValGlnArg 20

Qy   292 ATCAAGGCCTCGAAGTCATTCCGCACCCCGGCTACTCCACCACGACCCATGTTAATGAC 351
         |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    21 IleLysAlaSerLysSerPheArgHisProGlyTyrSerThrGlnThrHisValAsnAsp 40

Qy   352 CTCATGCTCGTGAAGCTCAATAGCCAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGG 411
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    41 LeuMetLeuValLysLeuAsnSerGlnAlaArgLeuSerSerMetValLysLysValArg 60

Qy   412 CTGCCCTCCGGTCGAGAGGTCCCTGGAACCCCTGACCACCTGTCTCCGGCTGGGCACTACC 471
         |||||||||||||||::|||||||||||||||||||||||||||||||
Db    61 LeuProSerArgCysGluProProGlyThrThrCysThrValSerGlyTrpGlyThrThr 80

Qy   472 ACGAGCCCAGATGTGACCTTTCCCTCTGCACCTCATGTGCTGGATGTCAAGCTCATCTCC 531
         |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    81 ThrSerProAspValThrPhaProSerAspLeuMetCysValAspValLysLeuIleSer 100

Qy   532 CCCAGGACTGCACGAAGGTTACAAGGACTTACTGGAAAATTCCATGCTGTGGCGCTGGC 591
         ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   101 ProGlnAspCysThrLysValTyrLysAspLeuLeuGluAsnSerMetLeuCysAlaGly 120

Qy   592 ATCCCCGACTCCAAGAAAAACGCCTGCAATGGTCAATGGTGACTCAGGGGACGGTTGGTGTGCAGA 651
         ||||||||||||||||||||||||||||||||||||||||||||||||
Db   121 IleProAspSerLysLysAsnAlaCysAsnGlyAspSerGlyGlyProLeuValCysArg 140

Qy   652 GGTACCCTGCAAGGTCTGGTGTCCTGGGAACTTTCCCTGCGGCCAACCCAATGACCCA 711
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   141 GlyThrLeuGlnGlyLeuValSerTrpGlyThrProPheProCysGlyThrAsnAspPro 160

Qy   712 GGAGTCTACACTCAAGTGTGCAAGTTCACCAAGTGGATAAATGACACCATGAAAAAGCAT 771
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   161 GlyValTyrThrGlnValCysLysPheThrLysTrpIleAsnAspThrMetLysLysHis 180

Qy   772 CGC 774
         |||
Db   181 Arg 181
```

```
Qy   331 ACACAGACCCATGTTAATGACCTCATGCTCGTGAAGCTCAATAGCCAGGCCAGGCTGTCA 390
         |||:||:||||||||||||||||||||::::||||||
Db   102 ThrLysThrHisValAsnAspIleMetLeuValArgLeuAspGluProValLysMetSer 121

Qy   391 TCCATGGTGAAGAAAGTCAGGCTGCCCTCCGGTCGAGAGGTCCCTGGAACCACCTGTACT 450
         ||| ||:|| ||:::|| |||||||
Db   122 SerLysValGluAlaValGlnLeuProGluHisCysGluProGlyThrSerCysThr 141

Qy   451 GTCTCCGGCTGGGCACTACCACCAGCAGCTTTCCCTCTGCACCTCATGTGC 510
         |||||||||||||||||||||||||||||||
Db   142 ValSerGlyTrpGlyThrThrThrSerProAspValThrPheProSerAspLeuMetCys 161

Qy   511 GTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTACAAGGACTTACTGGAA 570
         ||||||||||||||||   ||||||||||||
Db   162 ValAspValLysLeuIleSerProArgGluCysGlyLysLysValTyrLysAspLeuGly 181

Qy   571 AATTCCATGCTGTGGCGCTGGCATCCCGACTCCAAGAAAAACGCCTGCAATGGTGACTCA 630
         :::|||||||||||||     |||  ||||||||||||||
Db   182 LysThrMetLeuCysAlaGlyIleProAspSerLysThrAsnThrCysAsnGlyAspSer 201

Qy   631 GGGGACCCTGGTGTGCGAGAGGTACCCTGCAAGGTCTGGTGTCCTGGGGAACTTTCCCT 690
         ||||||||||||||  |||||||||||::||||
Db   202 GlyGlyProLeuValCysAsnAspProGlyLeuValSerGlnAspThrArgLeuTyrPro 221

Qy   691 TGCGGCCAACCCAATGACCCCAGGAGTCTACACTCAAGTGTGCAAGTTCACCAAGTGGATA 750
         ||||||||||||||:::  ||:|||||||
Db   222 CysGlyGlnProAsnAspProGlyValTyrThrGlnValCysLysTyrLysArgTrpVal 241

Qy   751 AATGACACCATGAAAAAGCATCGC 774
         ::|||||||| |||||
Db   242 MetGluThrMetLysThrHisArg 249
```

RESULT 4
Q8NFV7
ID   Q8NFV7   PRELIMINARY;   PRT;   181 AA.
AC   Q8NFV7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Kallikrein 7 short variant protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovarian carcinoma;
RX   MEDLINE=22623266; PubMed=12738725;
RA   Dong Y., Kaushal A., Brattsand M., Nicklin J., Clements J.A.;
RT   "Differential splicing of KLK5 and KLK7 in epithelial ovarian cancer
RT   produces novel variants with potential as cancer biomarkers.";
RL   Clin. Cancer Res. 9:1710-1720(2003).
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; AF411215; AAN03663.1; -.
DR   HSSP; P00760; 1EZX.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE 181 AA; 19987 MW; 86A28A03B80C2D78 CRC64;

Alignment Scores:
Pred. No.:
Score: 3.08e-77
             988.00

RESULT 5
Q9R048
ID   Q9R048   PRELIMINARY;   PRT;   234 AA.
AC   Q9R048;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Stratum corneum chymotryptic enzyme (Fragment).
GN   Name=Klk7;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL;
RX   MEDLINE=99399282; PubMed=10469296;
RA   Backman A., Stranden P., Brattsand M., Hansson L., Egelrud T.;
RT   "Molecular cloning and tissue expression of the murine analog to human
RT   stratum corneum chymotryptic enzyme.";
RL   J. Invest. Dermatol. 113:152-155(1999).
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; AF12429?; AAF01133.1; -.
DR   HSSP; P00760; 1EZX.
DR   MEROPS; S01.300; -.
DR   MGD; MGI:1346336; Klk7.

```
Qy   772 CGC 774
         |||
Db   181 Arg 181

Length:  181
Matches: 181
```

```
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
FT   NON_TER      234    234
SQ   SEQUENCE   234 AA;  25220 MW;  BEBE57D6C325B85F CRC64;
```

```
Alignment Scores:
Pred. No.:                    1
Score:                  4.89e-74            951.50
Percent Similarity:     84.98%
Best Local Similarity:  75.11%
Query Match:            53.46%
DB:                     2
```

```
NRPN_MOUSE
ID   NRPN_MOUSE     STANDARD;     PRT;     260 AA.
AC   Q61955;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Neuropsin precursor (EC 3.4.21.-) (NP) (Kallikrein 8).
GN   Name=Klk8; Synonyms=Nrpn, Prss19;
OS   Mus musculus (Mouse)
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Hippocampus;
RX   MEDLINE=95348817; PubMed=7621137;
RA   Chen Z.-L., Yoshida S., Kato K., Momota Y., Suzuki J., Tanaka T.,
RA   Ito J., Nishino H., Aimoto S., Kiyama H., Shiosaka S.;
RT   "Expression and activity-dependent changes of a novel limbic-serine
RT   protease gene in the hippocampus.";
RL   J. Neurosci. 15:5088-5097(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Yoshida S., Hirata A., Inoue N., Shiosaka S.;
RT   "Cloning and assignment of mouse neuropsin gene, Prss19 to chromosome
RT   7B4.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Czech II; TISSUE=Mammary gland;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [4]
RP   SEQUENCE OF N-TERMINUS, CHARACTERIZATION, AND MASS SPECTROMETRY.
RC   STRAIN=BALB/c; TISSUE=Brain;
RX   MEDLINE=98225302; PubMed=9556608; DOI=10.1074/jbc.273.18.11189;
RA   Shimizu C., Yoshida S., Shibata M., Kato K., Momota Y., Matsumoto K.,
RA   Shiosaka S., Midorikawa R., Kamachi T., Kawabe A., Shiosaka S.;
RT   "Characterization of recombinant and brain neuropsin, a plasticity-
RT   related serine protease.";
RL   J. Biol. Chem. 273:11189-11196(1998).
RN   [5]
RP   X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 33-257.
RC   TISSUE=Hippocampus;
RX   MEDLINE=99134351; PubMed=9933620; DOI=10.1074/jbc.274.7.4220;
RA   Kishi T., Kato M., Shimizu T., Kato K., Matsumoto K., Yoshida S.,
RA   Shiosaka S., Hakoshima T.;
RT   "Crystal structure of neuropsin, a hippocampal protease involved in
RT   kindling epileptogenesis.";
RL   J. Biol. Chem. 274:4220-4224(1999).
CC   -!- FUNCTION: Suggested to be involved in kindling epileptogenesis and
CC       hippocampal plasticity. Has a strong proteolytic activity against
CC       fibronectin.
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.
```

```
US-09-905-083A-30 (1-969) x Q9R048 (1-234)

Qy     31  CTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACTCAGAGGAAGAAGCC  90
               |||||  |||                ::|||||||||||||||||||||||||
Db      5  LeuLeuSerLeuIleThrValLeuLeuSerLeuAlaLeuGluThrAlaGly--------  21

Qy     91  CAGGGTGACAAGATTAATTGATGGCGCCCCATGTGCAAGAGGTTCCCACCCATGGCAGTG 150
               ||||||||        |||||::::||||||||||||||||||||||||||||||
Db     22  GlnGlyGluArgIleIleAspGlyTyrLysCysLysGluGlySerHisProTrpGlnVal  41

Qy    151  GCCCTGCTCCAGTGGCAATCAGCTCGCGAGGAGGCGTCCTGGTCAATGAGCGCTGGGTG 210
               ||||||||||||||||||||||||||||||||||||::::|| ||||||||||||
Db     42  AlaLeuLeuLysGlyAsnGlnLeuHisCysGlyGlyValLeuValAspLysTyrTrpVal  61

Qy    211  CTCACTGCGCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACGCTG 270
               ||||||||||||||||     ::||||  ||||||||| ||||||||  :::|||
Db     62  LeuThrAlaAlaHisCysLysLysMetGlyGlnTyrGlnValGlnLeuGlySerAspLysIle  81

Qy    271  GGGGACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGACCACCCCGGCTACTCC 330
               |||||:: |||||  |||||||||||||||||:::||||||||||||||||||||
Db     82  GlyAspGlnSerAlaGlnArgIleLysAlaSerLysSerPheArgHisProGlyTyrSer 101

Qy    331  ACACAGACCCCATGTTAATGACCTCGTGAAGCTCAATAGCCAGGCCAGGCTGTCA 390
               ||::|||||||||||  ||||||||||||||:::||  ::::: |||||||
Db    102  ThrLysThrHisValAsnAspIleMetLeuValArgLeuAspGluProValIleLysMetSer 121

Qy    391  TCCATGGTGAAGAAAGTCAGGCTGCCCTCGCCTGCGAACCACCTGTACT 450
               ||| ||::: ||:||||           |||||||||||||||||||||
Db    122  SerLysValGluAlaValGlnLeuProGluHisCysGluProProGlyThrThrSerCysThr 141

Qy    451  GTCTCCGGCTGTGGGCACTACCACGAGCCCAGATGTGACCTTTCCCTCTGACCTCATGTGC 510
               ||||||||||||||||||| |||||||||||||||||||||||||||
Db    142  ValSerGlyTrpGlyThrThrHisGlyProAspArgValThrPheProSerAspLeuMetCys 161

Qy    511  GTTGGATGTCAAGTCAGTTCATCTCCCCCCAGGACTGCACGAAGGTTTACAAGGACTTACTCGAA 570
               ||||||||||      :::|||        ||||||||  |||||||||||
Db    162  SerAspValLysSerLeuIleSerSerArgGluCysLysValTyrLysAspLeuLeuGly 181

Qy    571  AATTCCATGGTGCGCTGGCATCCCGACTCCAAGAGAAAAACGCCTGCAATGGTGACTCA 630
               ::::|||||||||||||||||||| |||          |||||||||||||||
Db    182  LysThrMetLeuCysAlaGlyIleProAspSerLysLysThrAsnThrCysAsnGlyAspSer 201

Qy    631  GGGGACCGTTGGTGCAGAGGTACCCTGCCAAGGTCTGGTGTCCTGGGGAACTTTCCCT 690
               ||||||||||||||||||||||||||||||||||||||||||||||||
Db    202  GlyGlyProLeuValCysAsnAspThrLeuGlnGlyLeuAlaSerArgGlyThrThrTyrPro 221

Qy    691  TGCGGCCACCCAATGACCCAGGAGTCTTCACACTCCAAGTG 729
               :::||||||||||||||||||||||||||||||||||||
Db    222  CysGlyGlnProAsnAspProGlyValTyrThrGlnVal 234
```

RESULT 6

```
CC   -!- ENZYME REGULATION: Strongly inhibited by diisopropyl
CC       fluorophosphate, leupeptin and (4-amidinophenyl)methanesulfonyl 1-
CC       fluoride.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed specifically in the limbic system of
CC       mouse brain and is localized at highest concentration in pyramidal
CC       neurons of the hippocampal CA1-3 subfields.
CC   -!- MASS SPECTROMETRY: MW=26613; METHOD=MALDI; RANGE=29-260;
CC       NOTE=Ref.4.
CC   -!- MASS SPECTROMETRY: MW=26229; METHOD=MALDI; RANGE=33-260;
CC       NOTE=Ref.4.
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
CC       subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; D30785; BAA06451.1; -.
DR   EMBL; AB032202; BAA92435.1; -.
DR   EMBL; BC055895; AAH55895.1; -.
DR   PIR; I56559; I56559.
DR   PDB; 1NPM; X-ray; A/B=33-256.
DR   MEROPS; S01.244; -.
DR   MGD; MGI:892018; Klk8.
DR   InterPro; IPR009903; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   3D-structure; Direct protein sequencing; Glycoprotein; Hydrolase;
KW   Serine protease; Signal; Zymogen.
FT   SIGNAL        1     28       Potential.
FT   PROPEP       29     32
FT   CHAIN        33    260
FT   ACT_SITE     73     73       Neuropsin.
FT   ACT_SITE    120    120       Charge relay system (By similarity).
FT   ACT_SITE    212    212       Charge relay system (By similarity).
FT   DISULFID     39    173       Charge relay system (By similarity).
FT   DISULFID     58     74
FT   DISULFID    145    246
FT   DISULFID    152    218
FT   DISULFID    184    198
FT   DISULFID    208    233
FT   CARBOHYD    110    110       N-linked (GlcNAc...) (Potential).
FT   STRAND       34     34
FT   STRAND       37     38
FT   TURN         41     42
FT   TURN         45     46
FT   STRAND       47     52
FT   TURN         53     54
FT   STRAND       55     64
FT   TURN         65     66
FT   STRAND       67     70
FT   HELIX        72     74
FT   STRAND       80     83
FT   STRAND       87     87
FT   STRAND       88     89
FT   TURN         89     89
FT   STRAND       96     98
FT   STRAND      100    105
FT   TURN        107    108
FT   TURN        114    115
FT   TURN        118    119
FT   STRAND      122    126
FT   STRAND      140    141
FT   TURN        148    149
FT   STRAND      151    156
FT   STRAND      170    170
FT   STRAND      172    178
FT   HELIX       181    187
FT   TURN        189    191
FT   TURN        194    195
FT   STRAND      196    200
FT   TURN        202    203
FT   STRAND      206    206
FT   TURN        209    210
FT   STRAND      212    213
FT   TURN        215    218
FT   TURN        219    220
FT   STRAND      221    228
FT   STRAND      235    235
FT   TURN        236    237
FT   STRAND      238    238
FT   STRAND      240    244
FT   HELIX       245    256
SQ   SEQUENCE   260 AA;  28523 MW;  BE5F6F6BE37CD60B CRC64;

Alignment Scores:
Pred. No.:                 1
Score:                  1.2e-44       Length:           260
Percent Similarity:    614.00         Matches:          114
Best Local Similarity: 61.94%         Conservative:      39
Query Match:           46.15%         Mismatches:        88
DB:                    34.49%         Indels:             6
                       1              Gaps:               3

US-09-905-083A-30 (1-969) x NRPN_MOUSE (1-260)

Qy     37 CCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAGAAGCCAGGGT  96
                 ::|||||||     ::|||||||||||||||||||||||
Db     11 ProTrpIleLeuLeuLeuPheMetGlyAlaTrpAlaGlyLeuThrArgAlaGlnGly  30

Qy     97 GACAAGATTATTGATGGCGCCCCCATGTGCAAGAGAGCTCCCACCCATGGCAGGTGGCCCTG 156
          ||||||:::::|||  ||       ||| |||||||||||||||||||||
Db     31 SerLysIleLeuGluGlyArgGluCysIleProHisSerGlnProTrpGlnAlaAlaLeu  50

Qy    157 CTCAGTGCAATCAGCTCCACTGCGGAGCGTCCTGGTCAATGAGCGCTGGGTGCTCACT 216
                 :::|||||       ||||||:||||       :::|||||||||||
Db     51 PheGlnGlyGluArgLeuIleCysGlyGlyValLeuValGlyAspArgTrpValLeuThr  70

Qy    217 GCCGCCCACTGCAAGATGAATGAATGACACCGTGCACCTGGGCAGTGATACGCTG------ 270
          ||||||||||     :::||:::||||||      :::||
Db     71 AlaAlaHisCysLysLysGlnLysTyrSerValArgLeuGlyAspHisSerLeuGlnSer  90

Qy    271 GGGCACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCCCACCCCGGCTACTCC 330
          |||::  ||||||||||       ||:::||| :::||||||
Db     91 ArgAspGlnProGluGlnGlnLeuGlnValAlaGlnSerIleGlnHisHisProCysTyrAsn 110

Qy    331 ACA-------CAGACCCATGTTAATGACCTCGTGAAGCTCAATAGCCAGGCC 381
                       :::|||||||||||||
Db    111 AsnSerAsnProGluAspHisSerHisAspIleMetLeuIleArgLeuGlnAsnSerAla 130

Qy    382 AGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTGCCTGCGAACCCCCTGGAACC 441
          ||||||||::::|||  |||  :::|||       |||
Db    131 AsnLeuGlyAspLysValLysProValGlnLeuAsnLeuAlaAsnLeuCysProLysValGlyGln 150

Qy    442 ACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCCCTCGAC 501
          ||||||       :::||||||||||||
Db    151 LysCysIleIleSerGlyTrpGlyTrpValThrSerProGlnGluAsnProAsnThr 170

Qy    502 CTCATGTGCGTGATGTCAAGCTCATCTCCCCAGGACTGCACGAAGGTTTACAAGGAC 561
          ||||||:::|||||||     |||:::||||||
Db    171 LeuAsnCysAlaGluValLysIleTyrSerGlnAsnLysCysGluArgAlaTyrProGly 190

Qy    562 TTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCAAGAAAAACGCTGCAAT 621
          |||       :::||:::|||||||
Db    191 LysIleThrGluGlyMetValCysAlaGly---SerSerAsnGlyAlaAspThrCysGln 209

Qy    622 GGTGACTCAGGGGACCGGTTGGTGTGCAGAGGTACCCTGCAGAGGTCTCGGGA 681
          ||||       |||  :::|||||||
```

```
Db     210 GlyAspSerGlyGlyProLeuValCysAspGlyMetLeuGlnGlyIleThrSerTrpGly 229
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Qy     682 ACTTTCCCTTGCGCCACCAATGACCCAGGAGTCTACACTCAGTCTGCAAGTTCACC 741
           :::   |||||||||:::|||         |||||||||||||||:::|||:::|||
Db     230 SerAspProCysGlyLysProGluLysProGlyValTyrThrLysIleCysArgTyrThr 249
               |||||||
Qy     742 AAGTGGATAAATGACACACATG 762
               |||||||
Db     250 ThrTrpIleLysLysThrMet 256
```

RESULT 7
ID   KLK5_HUMAN   STANDARD;   PRT;   293 AA.
AC   Q9Y337, Q9HBG8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Kallikrein 5 precursor (EC 3.4.21.-) (Stratum corneum tryptic enzyme)
DE   (Kallikrein-like protein 2) (KLK-L2) (UNQ570/PRO1132).
GN   Name=KLK5; Synonyms=SCTE;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Stratum corneum;
RX   MEDLINE=99445563; PubMed=10514489; DOI=10.1074/jbc.274.42.30033;
RA   Brattsand M., Egelrud T.;
RT   "Purification, molecular cloning, and expression of a human stratum
RT   corneum trypsin-like serine protease with possible function in
RT   desquamation.";
RL   J. Biol. Chem. 274:30033-30040(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20118156; PubMed=10652563;
RA   Yousef G.M., Luo L.-Y., Diamandis E.P.;
RT   "Identification of novel human kallikrein-like genes on chromosome
RT   19q13.3-q13.4.";
RL   Anticancer Res. 19:2843-2852(1999).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   PubMed=11054574; DOI=10.1016/S0378-1119(00)00382-6;
RA   Gan L., Lee I., Smith R., Argonza-Barrett R., Lei H., McCuaig J.,
RA   Moss P., Paeper B., Wang K.;
RT   "Sequencing and expression analysis of the serine protease gene
RT   cluster located in chromosome 19q13 region.";
RL   Gene 257:119-130(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA   Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA   Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA   Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
RA   Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA   Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA   Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA   Vandlen R., Watanabe C., Wieand D., Woods K., Xie M.-H., Yansura D.,
RA   Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA   Godowski P., Gray A.;
RT   "The secreted protein discovery initiative (SPDI), a large-scale
RT   effort to identify novel human secreted and transmembrane proteins: a
RT   bioinformatics assessment.";
RL   Genome Res. 13:2265-2270(2003).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovary;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: May be involved in desquamation.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed in skin, breast, brain and testis.
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
CC   subfamily.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF168768; AAF03101.1; -.
DR   EMBL; AF135028; AAD26429.1; -.
DR   EMBL; AF243527; AAG33358.1; -.
DR   EMBL; AY359010; AAQ89369.1; -.
DR   EMBL; BC008036; AAH08036.1; -.
DR   HSSP; P00760; 1EZX.
DR   MEROPS; S01.017; -.
DR   Genew; HGNC:6366; KLK5.
DR   MIM; 605643; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR   GO; GO:0005544; P:epidermal differentiation; TAS.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Signal.
FT   SIGNAL   1   22   Potential.
FT   CHAIN   23   293   Kallikrein 5.
FT   ACT_SITE   108   108   Charge relay system (By similarity).
FT   ACT_SITE   153   153   Charge relay system (By similarity).
FT   ACT_SITE   245   245   Charge relay system (By similarity).
FT   DISULFID   73   206   By similarity.
FT   DISULFID   93   109   By similarity.
FT   DISULFID   178   279   By similarity.
FT   DISULFID   185   251   By similarity.
FT   DISULFID   217   231   By similarity.
FT   DISULFID   241   266   By similarity.
FT   CARBOHYD   69   69   N-linked (GlcNAc...) (Potential).
FT   CARBOHYD   173   173   N-linked (GlcNAc...) (Potential).
FT   CARBOHYD   208   208   N-linked (GlcNAc...) (Potential).
FT   CARBOHYD   252   252   N-linked (GlcNAc...) (Potential).
FT   CONFLICT   25   56   Missing (in Ref. 3).
SQ   SEQUENCE   293 AA; 32020 MW; D92C92F5609E5946 CRC64;

Alignment Scores:
Pred. No.:   293
Score: 1.66e-44   Length: 293
                                  Matches: 114
Percent Similarity: 65.97%   Conservative: 43
                    612.50%
```

```
Best Local Similarity:    47.90%        Mismatches:    72
Query Match:              34.41%        Indels:         9
DB:                            1        Gaps:           4

US-09-905-083A-30 (1-969) x KLK5_HUMAN (1-293)

Qy   76 GCAGGAGAAGAAGCCCAGGGTGAC----------AAGATTATTGATGGCGCCCCATGT 123
           |||||||||||::|||:.||                ::|||||||::::.  |||
Db   54 AlaGlyGluAspAlaArgSerAspAspSerSerArgIleIleAsnGlySerAspCys 73

Qy  124 GCAAGAGGCTCCACCTGGGCAGTGGCC---CTGCTCAGTGCAATCAGCTCCACTGC 180
           :::  |||||||| ||||||||        ||||||  ||||||||||::::||
Db   74 AspMetHisThrGlnProTrpGlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCys 93

Qy  181 GGAGGGCGTCCTGGTCAATGAGCGCTGGCTGCTCACTGCCGCCCACTGCAAGATGAATGAG 240
           || |||||||::  ||:|:: :|||||||||||:::||||||||||||:||
Db   94 GlyAlaValLeuValHisProGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysVal 113

Qy  241 TACACCGTGCACCTGGGCAGTGATACGTGGGC--------GACAGGAGAGCTCAGAGG 291
           :::  ||||||||| ||||||||||||| |||        ||||||||||
Db  114 PheArgValArgLeuGlyHisTyrSerLeuSerProValTyrGluSerGlyGlnGlnMet 133

Qy  292 ATCAAGGCTCGAGTCATTCCGCACCCCGGCTACTCCACAGACCCATGTTAATGAC 351
           :::       |||||||                ||||||||||||
Db  134 PheGlnGlyValLysSerIleProHisProGlyTyrSerHisProGlyHisSerAsnAsp 153

Qy  352 CTCATGCTCGTGAAGCTCAATAGCCAGGCCAGGGTCGTCATCCATGGTGAAGAAAGTCAGG 411
           |||||||||||||||||| |||||        ::: |||:.|||:: :|| :::
Db  154 LeuMetLeuIleLysAsnArgArgIleArgProThrLysAspAspArgProIleAsn 173

Qy  412 CTGCCCTCCGGCTGCGAACCCCCTGGAACCACCTGTCTTCCGGGTGGGCACTACC 471
           :::  |||  |||||||||||||||||||||||||||||||||||||
Db  174 ValSerSerHisCysProSerAlaGlyThrLysCysLeuValSerGlyTrpGlyThrThr 193

Qy  472 ACGAGACCCAGATGTGACCTTCCCTGACCTCCATGTGCGTGAGTCAAGCTCATCTCC 531
           ||||||||||||||||||||||||  :::::::::   ::|||||||||||||||
Db  194 LysSerProGlnValHisProLysValLeuGlnCysLeuAsnIleSerValLeuSer 213

Qy  532 CCCCAGGAGACTGCACGAAGGTTTACAAGGACTTACTGGAAAATTCCATGCTGTCGCGCTGGC 591
           :::  ||| |||  ||:             ::::::::||   |||||||||||||
Db  214 GlnLysArgCysGluAspAlaTyrProArgGlnIleAspAspThrMetPheCysAlaGly 233

Qy  592 ATCCCCGACTCCAAGAAAAACGCCTGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGA 651
           :::|||             |||||||||||||||||||||||||||||||
Db  234 ---AspLysAlaGlyAlaArgAspSerCysGlnGlyAspSerGlyGlyProValValCysAsn 252

Qy  652 GGTACCCTGCAAGGTCTGGTGTCCTCGGGAACTTCCCTGCGCCAACCCAATGACCCA 711
           ||::|||||||||||||||||            ::|||||||||||
Db  253 GlySerLeuGlnGlyLeuValSerTrpGlyAspTyrProCysAlaArgProAsnArgPro 272

Qy  712 GGAGTCTTCACACTCCAAGTGTGCAAGTTCACCAAGTGGATAAATGACACCATGAAA 765
           |||||||||||||  ::|||||||||||||||||||||  :::|||:::
Db  273 GlyValTyrThrArgLeuCysLysPheThrLysTrpIleGlnThrIleGln 290
```

RESULT 8
```
NRPN_RAT
ID   NRPN_RAT    STANDARD;      PRT;     260 AA.
AC   O88780;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Neuropsin precursor (EC 3.4.21.-) (NP) (Kallikrein 8) (Brain serine
DE   protease 1).
GN   Name=Klk8; Synonyms=Bsp1, Nrpn, Prss19;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Fischer; TISSUE=Brain;
RX   MEDLINE=98389725; PubMed=9722524; DOI=10.1074/jbc.273.36.23004;
RA   Davies B.J., Pickard B.S., Steel M., Morris R.G.M., Lathe R.;
RT   "Serine proteases in rodent hippocampus.";
RL   J. Biol. Chem. 273:23004-23011(1998).
CC   -!- FUNCTION: Suggested to be involved in kindling epileptogenesis and
CC       hippocampal plasticity. Has a strong proteolytic activity against
CC       fibronectin (By similarity).
CC   -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-, Lys-.
CC   -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC   -!- TISSUE SPECIFICITY: Restricted to hippocampus.
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
CC       subfamily.
CC   --------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------------------
DR   EMBL; AJ005641; CAA06643.1; -.
DR   HSSP; Q61955; 1NPM.
DR   MEROPS; S01.244; -.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Glycoprotein; Hydrolase; Serine protease; Signal; Zymogen.
FT   SIGNAL        1     28       Potential.
FT   PROPEP       29     32       By similarity.
FT   CHAIN        33    260       Neuropsin.
FT   ACT_SITE     73     73       Charge relay system (By similarity).
FT   ACT_SITE    120    120       Charge relay system (By similarity).
FT   ACT_SITE    212    212       Charge relay system (By similarity).
FT   DISULFID     39    173       By similarity.
FT   DISULFID     58     74       By similarity.
FT   DISULFID    145    246       By similarity.
FT   DISULFID    152    218       By similarity.
FT   DISULFID    184    198       By similarity.
FT   DISULFID    208    233       By similarity.
FT   CARBOHYD    110    110       N-linked (GlcNAc...). (Potential).
SQ   SEQUENCE   260 AA;  28510 MW;  58DF4F0602A0B7F5 CRC64;

Alignment Scores:
Pred. No.:                              Length:               260
Score:                   611.50         Matches:              114
Percent Similarity:      62.35%         Conservative:          40
Best Local Similarity:   46.15%         Mismatches:            86
Query Match:             34.35%         Indels:                 7
DB:                           1         Gaps:                   4

US-09-905-083A-30 (1-969) x NRPN_RAT (1-260)

Qy   46 ATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGA---GAAGAAGCCCAGGGTGACAAG 102
           ||||:|||||                               :||  ||  ::||||
Db   13 IleLeuLeuPheLeuLeuMetGlyAlaTrpAlaGlyLeuThrArgAlaGlnGlySerLys 32

Qy  103 ATTATTGATGGCGCCCCATGTCAAGAGGCTCCCACCATGGCAGGTGGCCCCATGCTCAGT 162
           ||:::::||:||                |||| |||||||||||||:||
Db   33 IleLeuGluGlyGlnGluCysLysProHisSerGlnProHisProTrpProAlaLeuPheGln 52

Qy  163 GGCAATCAGCTCCACTGCGGAGGGCGTCCTGGTCAATGAGCGCTGGCTGCTCACTGCCGCC 222
           ||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db   53 GlyGluArgLeuValCysGlyGlyValLeuValGlyAspTrpAsnGlyArgThrAlaAla 72

Qy  223 CACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGTGATACGTG------GGCGAC 276
           |||||||||||| |||                   :::||||||:||
Db   73 HisCysLysLysAspLysTyrSerValArgGlyLeuGlyLeuGlnLysArgAsp 92
```

```
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   276 AA;  30360 MW;  33E189C67492FDCA CRC64;

Alignment Scores:
Pred. No.:                              3e-44          Length:              276
Score;                                  609.50         Matches:             120
Percent Similarity:                     63.18%         Conservative:         43
Best Local Similarity:                  46.51%         Mismatches:           76
Query Match:                            34.24%         Indels:               19
DB:                                     2              Gaps:                  6

US-09-905-083A-30 (1-969) x Q8CGR6 (1-276)
```

Qy   277 AGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGCCACCCGGCTACTCCACA--- 333
         ||| ||::::||| :::::||||| :::||||
Db   93 GlyProGluGluGlnValAlaArgSerIleGlnHisProCysPheAsnSerSer   112

Qy   334 ------CAGACCCCATGTTAATGACCTCATGCTCGTCGTGAAGCTCAATAGCCAGGCCAGGCTG 387
               ::: ||| ||| :::
Db   113 AsnProGluAspHisSerHisAspIleMetLeuIleArgLeuGlnAsnSerAlaAsnLeu   132

Qy   388 TCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCGCTCGGAACCCCTGGAACCACCTGT   447
         || ||| || ||| |||
Db   133 GlyAspLysValLysProIleGluLeuAlaAsnLeuCysLeuGlyProLysValGlyGlnLysCys 152

Qy   448 ACTGTCTCCGGCTGGGGCACTACCACCAGAGCCCAGATGTGACCTTTCCCTCTGACCTCATG   507
         |||||| |||||||
Db   153 IleIleSerGlyTrpGlyThrValThrSerProGlyGluAsnPheProAsnThrLeuAsn   172

Qy   508 TGCCTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTACAAGGACTTACTG   567
         || ::: |||||||::: ||| || ::: |||
Db   173 CysAlaGluValLysIleTyrSerGlnAsnLysCysGluArgGlaTyrProGlyLysIle   192

Qy   568 GAAAATTCCATGCTGTGCGCTGGCATCCCCGACTCCCAGACCCTGCAATGGTGAC   627
         |||||| ::: || ||||||
Db   193 ThrGluGlyMetValCysAlaGly--SerSerAsnGlyAlaAspThrCysGlnGlyAsp   211

Qy   628 TCAGGGGGACCCGTTGGTGTGCGAGAGGTACCCTGCCAAGTGTCGTGTCCTGGGGAACTTC   687
         ||||||||||||||| ::: |||||||| :::||||||
Db   212 SerGlyGlyProLeuValCysAsnGlyValLeuGlnGlyIleThrThrTrpGlySerAsp   231

Qy   688 CCTTGCCGGCCAACCAATGACCCAGGAGTCTACACTCAAGTGTCAAGTTCACCAAGTGG   747
         ||||||||||||||||||||| ||||||||::: ||||||| :::||
Db   232 ProCysGlyLysProGluLysProGlyValTyrThrLysIleCysArgTyrThrArgTrp   251

Qy   748 ATAAATGACCACATGAAAAAG 768
         ||| ||
Db   252 IleLysLysThrMetGlyLys 258

RESULT 9
Q8CGR6
ID   Q8CGR6   PRELIMINARY;   PRT;   276 AA.
AC   Q8CGR6;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Glandular kallikrein KLK13.
GN   Name=Klk13;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22325484; PubMed=12437987; DOI=10.1016/S0006-291X(02)02629-3;
RA   Olsson A.Y., Lundwall A.;
RT   "Organization and evolution of the glandular kallikrein locus in Mus
RT   musculus.";
RL   Biochem. Biophys. Res. Commun. 299:305-311(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Adams M., Mural R.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; AY152432; AAN78420.1; -.
DR   HSSP; P00760; 1BZX.
DR   MEROPS; S01.306; -.
DR   MGD; MGI:95292; Klk13.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
```

Qy   46 ATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAAGAGCCCAGGTGACAAGATT 105
          ||| |||||||||||||| ::: |||
Db   8 IleAlaCysTyrLeuThrLeuAlaLeu--SerGluGlyIleSerArgAspTyrProLysIle 26

Qy   106 ATTGATGGGCGC------------------------------------CCATGTGCAAGAGGC 132
          :::|||||
Db   27 LeuAsnGlyThrAsnGlyThrSerGlyPheLeuProGlyTyrThrCysLeuProHis 46

Qy   133 TCCCCCATGGCAGGTGCGCCTGTCCAGTGGCAATCAGCTCCACTGCGGAGGCGTCCTG 192
          ||| ||| :::||||||||||||
Db   47 SerGlnProTrpGlnAlaAlaAlaLeuLeuIleLeuArgGlyArgLeuLeuCysGlyGlyValLeu 66

Qy   193 GTCAATGAGCGCTGGGTGCTCACTGCCGCCCACTGCCAAGATGAATGAGTACACCGTGCAC 252
          :::|||||||||| ::: ||| ::: :::
Db   67 ValHisProLysTrpValLeuThrAlaAlaHisCysArgLysAspGlyTyrThrValHis 86

Qy   253 CTGGGCAGTGATACGCTGGC------------GACAGGAGAGCTCAGAGGATCAAGGCCTCGAAG 306
          ||||| ||||
Db   87 LeuGlyLysHisAlaLeuGlyArgValGluAsnGlyGluGlnAlaMetGluValValArg 106

Qy   307 TCATTCCGCCACCCCGGCTACTCC---ACACAGACCCCATGTTAAT--------GACCTC 354
          ||||| ::::|||
Db   107 SerIleProHisProGluTyrGlnValThrProThrHisLeuAsnHisAspHisAspIle 126

Qy   355 ATGCTCGTCGTGAAGCTCAATAGCCAGGCCAGGCTGCCCTCCCGCTCGTGAAGAAAGTCAGGCTG 414
          |||||::::::|||||::::::   ::::::|||||
Db   127 MetLeuLeuGluLysSerProValGlnLeuSerSerHisValArgLeuLysLeu 146

Qy   415 CCCTCC---CGCTGCGAACCCCCTGGAACCACCTGTACTGTCTCCGGCTGGGGCACTACC 471
          :::::: ||||||||||::::::|||||||||
Db   147 SerAlaAspAspCysLeuProThrGlyThrCysArgValSerGlyValSerGlyTrpGlyThrThr 166

Qy   472 ACGAGCCCAGATGTGACCTTTCCCTCTGACCTCATGTGCCTGGATGTCAAGCTCATCTCC 531
          ||||||||| ::: || ||||||||
Db   167 ThrSerProGlnValAsnTyrProLysThrLeuGlnCysAlaAsnIleGluLeuArgSer 186

Qy   532 CCCAGGACTGCACGAAGGTTACAAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGC 591
          :::::::||| :::::::||||||||
Db   187 AspGluGluCysArgGlnValTyrProGlyLysIleThrAlaAsnMetLeuCysAlaGly 206

Qy   592 ATCCCCGACTCCAAGAAAAACGCCTGCAATGGTGACCTGCAGGGGGACCGTTGGGTGTGCAGA 651
          ||||||||| ||||||| ::::::
Db   207 ThrLysGluGlyLysLysAspSerCysGluGlyGlyLysProLeuIleCysAsn 226

Qy   652 GGTACCCTGCAAGGTCTGTGTCCTGGGGAACTTCCCTGCCGGCCAACCAATGACCCA 711
          |||||||||||| ::::::::::::|||||
Db   227 GlyLysLeuTyrGlyIleIleSerTrpGlyAspProCysGlyLysProAsnArgPro 246

Qy   712 GGAGTCTACACTCAAGTGTCAAGTTCACCAAGTGGATAAATGACCACATGAAA 765
          |||||||||::::::||||||| :::||||||:::
Db   247 GlyValTyrThrArgValSerLysTyrLeuArgTrpIleLeuArgGluIleIleArg 264

RESULT 10
Q6P3Z0
ID   Q6P3Z0   PRELIMINARY;   PRT;   246 AA.
AC   Q6P3Z0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

```
Db        50 LysAspGlyTyrThrValHisLeuGlyLysHisAlaLeuGlyArgValGluAsnGlyGlu 69
             :::        ||||||||||||           :::           |||||
Qy       286 CAGAGGATCAAGGCTCGAAGTCATTCCGCCACCCGGCTACTCC---ACACAGACCCAT 342
             ||||||  ||||||||||||  :::::::|||||||||||||| |||  |||||||||||
Db        70 GlnAlaMetGluValValArgSerIleProHisProGluTyrGlnValThrProThrHis 89

Qy       343 GTTAAT--------GACCTCATGTCGTGAAGCTCAAATAGCCAGGCCAGGCTGTCATCC 393
             :::|||         ||:::|::::::::::::|||||||||||||| |||||||||||||
Db        90 LeuAsnHisAspHisAspIleMetLeuLeuGluLeuLysSerProValGlnLeuSerSer 109

Qy       394 ATGTGAAGAAAGTCAGGCTGCCCTCC---CGCTGCGAACCCCTGGAACCACCTGTACT 450
             ||||::  :::::::|||  :::  ||| ||||||:::  |||  :::|||||||||||
Db       110 HisValArgThrLeuLysLeuSerAlaAspAspCysLeuProThrGlyThrCysCysArg 129

Qy       451 GTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCCTCTGACCTCATGTGC 510
             |||  ||||||  |||||||||||||||||||||||||||  :::||| |||||||||||
Db       130 ValSerGlyTrpGlyThrThrThrSerProGlnValAsnTyrProLysThrLeuGlnCys 149

Qy       511 GTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTACAAGGACTTACTGGAA 570
             ||||||  :::::::|||         :::|||||||  :::|||||
Db       150 AlaAsnIleGluLeuArgSerAspGluArgValTyrProGlyLysIleThr 169

Qy       571 AATTCCATGCTGTGCGCTGGCATCCCCGACTCCAAGAAAAAGCCTGCAATGGTGACTCA 630
             :::|||||||||||  ||||||||  :::  ||||||||||||:::||| ||||||||||
Db       170 AlaAsnMetLeuCysAlaGlyThrLysGlyLysAspSerCysGluGlyAspSer 189

Qy       631 GGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTCTGGTGTCCTGGGAACTTTCCCT 690
             ||||||||||||||||:::|||       ||||||||||||||||||||||||||||||
Db       190 GlyGlyProLeuIleCysAsnGlyLysLeuTyrGlyIleIleSerTrpGlyAspPhePro 209

Qy       691 TGCGGCCAACCCAATGACCCAGGAGTCTACACTCCAAGTGCAAGTGCAAGTGGATA 750
             ||||||||||||||||  ||||||||||||||||||||||||||::  |||||||||
Db       210 CysGlyGlnProAsnArgProGlyValTyrThrArgValSerTyrLeuArgTrpIle 229

Qy       751 AATGACACCATGAAA 765
             :::  :::::
Db       230 ArgGluIleIleArg 234
```

RESULT 11
Q8CGR5
ID   Q8CGR5   PRELIMINARY;   PRT;   250 AA.
AC   Q8CGR5;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Glandular kallikrein KLK14.
GN   Name=Klk14;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=22325484; PubMed=12437987; DOI=10.1016/S0006-291X(02)02629-3;
RA   Olsson A.Y., Lundwall A.;
RT   "Organization and evolution of the glandular kallikrein locus in Mus
RT   musculus.";
RL   Biochem. Biophys. Res. Commun. 299:305-311(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Adams M., Mural R.;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; AY152433; AAN78421.1; -.
DR   HSSP; P00760; 1EZX.
DR   MEROPS; S01.029; -.
DR   MGD; MGI:2447564; Klk14.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.

DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Jaw and limb;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.B., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Jaw and limb;
RA   Strausberg R.;
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; BC063763; AAH63763.1; -.
DR   HSSP; P00761; 1AKS.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0008233; F:peptidase activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Hypothetical protein; Protease; Serine protease.
SQ   SEQUENCE   246 AA; 27213 MW; 0BC1454D6CC8CDEC CRC64;

Alignment Scores:
Pred. No.:
Score: 7.98e-44   Length: 246
                     604.50   Matches: 112
Percent Similarity: 67.11%   Conservative: 39
Best Local Similarity: 49.78%   Mismatches: 67
Query Match: 33.96%   Indels: 7
DB: 2   Gaps: 4

US-09-905-083A-30 (1-969) x Q6P3Z0 (1-246)

```
Qy       112 GGGCGCCCCATGTGCCAAGAGAGGCTCCCACCCATGGCAGGTGGCCCTGCTCAGTGGCAATCAG 171
             |||  |||       |||  |||||||||||||||||||||||||||
Db        10 GlyTyrThrCysLeuProHisSerGlnProTrpGlnAlaAlaLeuLeuIleArgGlyArg 29

Qy       172 CTCCACTGCGGAGGCGTCCTGGTCAATGAGCGCTGGGTCTCACTGCCCGCCCACTGCAAG 231
             |||  ||||||||| ||||||:::  :::|||||||||||||||||||||  :::||||||
Db        30 LeuLeuCysGlyGlyValLeuValHisProLysTrpValLeuThrAlaAlaHisCysArg 49

Qy       232 ATGAATGAGTACACCGTGCACCTGGGCAGTAGTAGCGTGATACGCTGGGC-----GACAGGAGAGCT 285
```

```
DR    InterPro; IPR001314; Peptidase_S1A.
DR    InterPro; IPR009003; Pept_Ser_Cys.
DR    Pfam; PF00089; Trypsin; 1.
DR    PRINTS; PR00722; CHYMOTRYPSIN.
DR    SMART; SM00020; Tryp_SPc; 1.
DR    PROSITE; PS50240; TRYPSIN_DOM; 1.
DR    PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR    PROSITE; PS00135; TRYPSIN_SER; 1.
KW    Hydrolase; Protease; Serine protease.
SQ    SEQUENCE  250 AA;  27016 MW;  F62FEBP2290FEBB8 CRC64;
```

Alignment Scores:

```
Pred. No.:                           Length:        250
Score:                 7.27e-43      Matches:       114
Percent Similarity:    593.50        Conservative:   42
Best Local Similarity: 61.66%        Mismatches:     86
Query Match:           45.06%        Indels:         11
DB:                    33.34%        Gaps:            4
```

RESULT 12
Q9D140
```
ID   Q9D140        PRELIMINARY;       PRT;      293 AA.
AC   Q9D140;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched
DE   library clone:1110030019 product:weakly similar to KALLIKREIN 5 (EC
DE   3.4.21.-) (STRATUM CORNEUM TRYPTIC ENZYME) (KALLIKREIN-LIKE PROTEIN 2)
DE   (KLK-L2).
GN   Name=1110030019Rik;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RX   MEDLINE=20310913; PubMed=11076861; DOI=10.1101/gr.152600;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka K., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Whole body;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono M.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Kondo M., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata Y., Shibata K., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
```

```
US-09-905-083A-30 (1-969) x Q8CGR5 (1-250)

Qy    28  CTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTGGAAACTGCAGGAGAAGAA  87
             ||||||:||  ||||||  |||      |||||:||:|||
Db     4  LeuLeuIleIleLeuGlnAlaLeuAlaLeuAlaValAlaIleAla----------Gln  17

Qy    88  GCCCAGGGTGAC---AAGATTATTGATGCGCCCCATGTGCAAGAGGTCCACCCATGG  144
             :::  ||||||       ||| ||||||||| |||||||||:|||||||||||
Db    18  SerGlnGlyAspHisLysIleIleGlyIleGlyTyrArgCysValArgAsnSerGlnProTrp  37

Qy   145  CAGGTGGCCCTGCTCAGT-----GGCAATCAGCTCCACTGCGGAGGGTCCTGGTCAAT  198
             |||||||||  :::       ||||||:||:||||
Db    38  GlnValAlaLeuGlnAlaGlyProGlyHisArgPheLeuCysGlyGlyValLeuLeuSer  57

Qy   199  GAGCGCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGC  258
             ::::||||||:|||||||::||||||||||      |||  ||||||
Db    58  AspGlnTrpValIleThrAlaAlaHisCysAlaArgProIleLeuHisValAlaLeuGly  77

Qy   259  AGTGATACGCTG------GGCGACAGGAGAGCTCAGAGGATCAAGGCCTGAAGTCATTC  312
                               ::  ||||||  :::::::
Db    78  LysHisAsnIleArgArgTrpGluAlaThrGlnGlnHisValValArgValAlaArgGlnVal  97

Qy   313  CGGCACCCCGGCTACTCCACACAGACACCCATGTTAATGACCTCGTGTGAAGCTCAAT  372
             ||||||||  ||||||||||||||||  |||  |||||||:::||||||
Db    98  ProHisProGlnProGlnAlaHisAlaHisAspAsnAspLeuMetLeuLeuLysLeuGln  117

Qy   373  AGCCAGGCCAGGCTGCTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCGCTGCGAACCC  432
             ::: |||||||  |||  :::    ::: |||  ||
Db   118  LysLysValArgArgLeuGlyArgAlaValLysThrIleSerValAlaSerSerCysAlaSer  137

Qy   433  CCTGGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACCAGAGCCCAGATGTGACCTTT  492
             ::: |||||| ||:||||||:|||||||||||   ||||||  ||:||
Db   138  ProGlyThrProCysSerArgValSerGlyTrpGlyThrIleAlaSerProIleAlaArgTyr  157

Qy   493  CCCTCTGACCTCATGTGCCTGGATGTCAAGCTCCATCTCCCCCCAGGACTGCACGAAGTT  552
             |||:::  ||||||        ::: |||||
Db   158  ProThrAlaLeuGlnCysValAsnValAsnIleMetSerGluGlnAlaCysHisArgAla  177

Qy   553  TACAAGGACTTACTGGAAAATTCCATGCTGCGCTGGCATCCCGACTCCAAGAAAAAC  612
             |||      :::  ||:||||||:||   ||:||||||
Db   178  TyrProGlyIleIleThrSerGlyMetValCysValProGluGlyGlyValLysLysAsp  197

Qy   613  GCCTGCAATGGTCAGGGGGACCGGTTGGTGTGCAGAGGTACCCTGCAAGTGTCTGTG  672
             |||||||||    ||||   ||:|||||||
Db   198  SerCysGlnGlyAspSerGlyGlyProLeuValCysGlyGlyGlnLeuGlyLeuVal  217

Qy   673  TCCTGGGAACTTTCCCTGCGGCAACCCAATGACCCAGGAGTCTCACTCAAGTGTGC  732
             |||||||          |||||||||||||:::||||||||||
Db   218  SerTrpTrpGlyMetGluArgCysAlaAsnMetProGlyTyrProGlyValTyrValPheTyrAlaAsnLeuCys  237

Qy   733  AAGTTCACCAAGTGGATAAAATGACACCATGAAAAAGCAT  771
             ::: ||||||     ||: :::
Db   238  AsnTyrHisSerTrpIleGlnArgThrMetGlnSerAsn  250
```

RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: Belongs to peptidase family S1.
DR EMBL; AK003996; BAB23113.1; -.
DR HSSP; P00760; 1EZX.
DR MEROPS; S01.418; -.
DR MGD; MGI:1915918; 1110030O19Rik.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR001254; Peptidase_S1.
DR InterPro; IPR001314; Peptidase_S1A.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR Pfam; PF00089; Trypsin; 1.
DR PRINTS; PR00722; CHYMOTRYPSIN.
DR SMART; SM00020; Tryp_SPc; 1.
DR PROSITE; PS50240; TRYPSIN_DOM; 1.
DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR PROSITE; PS00135; TRYPSIN_SER; 1.
KW Hydrolase; Protease; Serine protease.
SQ SEQUENCE 293 AA.; 31908 MW; ED1F45D8226FB911 CRC64;

Alignment Scores:
Pred. No.:   1.85e-42   Length: 293
Score: 589.00   Matches: 108
Percent Similarity: 63.60%   Conservative: 44
Best Local Similarity: 45.19%   Mismatches: 79
Query Match: 33.09%   Indels: 8
DB: Gaps: 4

US-09-905-083A-30 (1-969) x Q9D140 (1-293)

[sequence alignment blocks - Qy and Db lines with nucleotide/protein sequences and match bars]

```
Percent Similarity:      61.83%         Conservative:       38
Best Local Similarity:   46.06%         Mismatches:         87
Query Match:             32.95%         Indels:              5
DB:                                     Gaps:                3

US-09-905-083A-30 (1-969) x Q80VS4 (1-242)

Qy      64  GCCTTGGAAACTGCAGGAGAAGAGCCCAGGGTGAC---AAGATTATTGATGGCGCCCCA 120
                |||||| |||||||| | ||||||||||||||||    |||||||||||
Db       2  AlaLeuAlaValAlaIleAlaGlnSerGlnGlyAspHisLysIleIleGlyGlyTyrArg  21

Qy     121  TGTGCAAGAGGCTCCCACCCATGCAGGTGGCCCTGCTCAGT-----GGCAATCAGCTC 174
                |||  |||||||||||||||||        |||  :::      ||||||||||
Db      22  CysValArgAsnSerGlnProTrpGlnValAlaLeuGlnAlaGlyProGlyHisArgPhe  41

Qy     175  CACTGCGGAGGCGGCTCCTGGTCAATGAGCCGCTGGGTGCTCACTGCGCGCCACTGCAAGATG 234
                ||||||||||||||| |||||:::|||||||| ||||||||| || ::::||||||
Db      42  LeuCysGlyGlyValLeuLeuSerAspGlnTrpValIleThrAlaAlaHisCysAlaArg  61

Qy     235  AATGAGTACACCGTGCACCTGGGCAGTGATACCGCTG------GGCGACAGGAGAGCTCAG 288
                   ||||| ||||||| ||||||        ::     ||||||||||||||:::
Db      62  ProIleLeuHisValAlaLeuGlyLysHisAsnIleArgArgTrpGluAlaThrGlnGln  81

Qy     289  AGGATCAAGGCCTTGAAGTCATTCCGCCACCCCGGCTACTCCACACAGACCCATGTTAAT 348
                :::::  :::::::  ||||||    ||||||||  ||  ||||||||||  |||
Db      82  ValValArgValAlaArgGlnValProHisProGlnTyrGlnAlaHisAspAsn 101

Qy     349  GACCTCATGTCTGCTGTGAAGCTCAATAGCCAGGCAGGCTGTCATCCATGTGAAGAAAGTC 408
                |||||||||||||||      || |||||:::|||     |||  ||||||||:::
Db     102  AspLeuMetLeuLeuLysLeuGlnLysLysValArgLeuGlyArgAlaValLysThrIle  121

Qy     409  AGGCTGCCCTCCCCGGCTGGCGAACCCACCTGGTACGTCTCCGGCTGGGGCCACT 468
                ::  ||| ||||||||||||||||||||||||||||||||||||||||||
Db     122  SerValAlaSerSerCysAlaSerProGlyThrProCysArgValSerGlyValTrpGlyThr 141

Qy     469  ACCAGCAGCCAGATGCTGACCTTTCCCTCTGACCTCATGTGCGCGGATGTCAAGCTCATC 528
                :::||||||||||||||||| :::||   ||  |||  ::::::
Db     142  IleAlaSerProIleAlaArgTyrProThrAlaLeuGlnCysValAsnValAsnIleMet 161

Qy     529  TCCCCCCAGGACTGCACGAAGGTTACAAGGACTTACTGGAAAATTCCATGTGTGCGCT 588
                |||   ||| || |||  :::::  |||||||   ::||| ||||||||
Db     162  SerGluGlnAlaCysHisArgAlaTyrProGlyIleIleThrSerGlyMetValCysAla 181

Qy     589  GGCATCCCGACTCCAAGAAAACGCCTGCAATGGTGACTCAGGGGACCGTTGGTGTGC 648
                |||  :::||||::  ||||||||| |||||||||||||||||||||
Db     182  GlyValProGluGlyGlyLysAspSerCysGlnGlyAspSerGlyGlyProLeuValCys 201

Qy     649  AGAGGTACCCTGCAAGGTGCGGGTTCTGGTGTCCTGGGAAACTTTCCCTTGCGGCAACCAATGAC 708
                |||||| |||||||||||      :::|| | |||||||||||||||
Db     202  GlyGlyGlnLeuGlnGlyValSerTrpGlyValSerGlyMetGluArgCysAlaMetProGlyTyr 221

Qy     709  CCAGGAGTCTACACTCAAGTGTGCAAGTTCACCAAGTGGATAAATGACACCATGAAAAAG 768
                :::|| |||  ::: :::||||||||  |||| |||||||||||||
Db     222  ProGlyValTyrAlaAsnLeuCysAsnTyrHisSerTrpIleGlnArgThrMetGlnSer 241

Qy     769  CAT 771
            :::
Db     242  Asn 242
```

RESULT 14
ID   KLKE_HUMAN   STANDARD;   PRT;   251 AA.
AC   Q9P0G3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Kallikrein 14 precursor (EC 3.4.21.-) (Kallikrein-like protein 6)
DE   (KLK-L6).
GN   Name=KLK14; Synonyms=KLKL6;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;

RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yousef G.M., Diamandis E.P.;
RT   "Molecular characterization, mapping, and tissue expression of KLKL6,
RT   a hormonally regulated kallikrein-like gene.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX   MEDLINE=21250997; PubMed=11352573; DOI=10.1006/geno.2000.6490;
RA   Hooper J.D., Bui L.T., Rae F.K., Harvey T.J., Myers S.A.,
RA   Ashworth L.K., Clements J.A.;
RT   "Identification and characterization of KLK14, a novel kallikrein
RT   serine protease gene located on human chromosome 19q13.4 and expressed
RT   in prostate and skeletal muscle.";
RL   Genomics 73:117-122(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   PubMed=15057824; DOI=10.1038/nature02399;
RA   Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J.,
RA   Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Tran-Gyamfi M.,
RA   Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E.,
RA   Caenepeel S., Carrano A.V., Caoile C., Chan Y.M., Christensen M.,
RA   Cleland C.A., Copeland A., Dalin E., Dehal P., Denys M., Detter J.C.,
RA   Escobar J., Flowers D., Fotopulos D., Garcia C., Georgescu A.M.,
RA   Glavina T., Gomez M., Gonzales E., Groza M., Hammon N., Hawkins T.,
RA   Haydu L., Ho I., Huang W., Israni S., Jett J., Kadner K., Kimball H.,
RA   Kobayashi A., Larionov V., Leem S.-H., Lopez F., Lou Y., Lowry S.,
RA   Malfatti S., Martinez D., McCready P.M., Medina C., Morgan J.,
RA   Nelson K., Nolan M., Ovcharenko I., Pitluck S., Pollard M.,
RA   Popkie A.P., Predki P., Quan G., Ramirez L., Rash S., Retterer J.,
RA   Rodriguez A., Rogers S., Salamov A., Salazar A., She X., Smith D.,
RA   Slezak T., Solovyev V., Thayer N., Tice H., Tsai M., Ustaszewska A.,
RA   Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Dubchak I.,
RA   Furey T.S., DeJong P., Dickson M., Gordon D., Eichler E.E.,
RA   Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M.,
RA   Rubin E.M., Lucas S.M.;
RT   "The DNA sequence and biology of human chromosome 19.";
RL   Nature 428:529-535(2004).
RN   [4]
RP   TISSUE SPECIFICITY.
RX   MEDLINE=20545474; PubMed=10969073; DOI=10.1074/jbc.M004525200;
RA   Harvey T.J., Hooper J.D., Myers S.A., Stephenson S.A., Ashworth L.K.,
RA   Clements J.A.;
RT   "Tissue-specific expression patterns and fine mapping of the human
RT   kallikrein (KLK) locus on proximal 19q13.4.";
RL   J. Biol. Chem. 275:37397-37406(2000).
CC   -!- SUBCELLULAR LOCATION: Secreted (Probable).
CC   -!- TISSUE SPECIFICITY: High expression in brain, bone marrow and
CC   fetal liver. Also expressed in liver, pancreas, fetal spleen,
CC   prostate and skeletal muscle.
CC   -!- SIMILARITY: Belongs to the peptidase S1 family. Kallikrein
CC   subfamily.
CC   --
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --
DR   EMBL; AF161221; AAD50773.2; -.
DR   EMBL; AF283669; AAK48523.1; -.
DR   EMBL; AF283670; AAK48524.1; -.
DR   EMBL; AC011473; AAG23260.1; -.
DR   HSSP; P00760; 1EZX.
DR   MEROPS; S01.029; -.
DR   Genew; HGNC:6362; KLK14.
DR   MIM; 606135; -.
DR   GO; GO:0005576; C:extracellular; NAS.
DR   GO; GO:0004252; F:serine-type endopeptidase activity; NAS.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; NAS.
DR   InterPro; IPR009003; Pept_Ser_Cys.

```
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; 1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Serine protease; Signal; Zymogen.
FT   SIGNAL         1     18      Potential.
FT   PROPEP        19     24      Activation peptide (Potential).
FT   CHAIN         25    251      Kallikrein 14.
FT   ACT_SITE      67     67      Charge relay system (By similarity).
FT   ACT_SITE     111    111      Charge relay system (By similarity).
FT   ACT_SITE     204    204      Charge relay system (By similarity).
FT   DISULFID      31     52      By similarity.
FT   DISULFID      52     68      By similarity.
FT   DISULFID     143    210      By similarity.
FT   DISULFID     175    189      By similarity.
FT   DISULFID     200    225      By similarity.
SQ   SEQUENCE   251 AA;  27452 MW;  9087953BAFA7ED25 CRC64;
```

```
Alignment Scores:
Pred. No.:             5.97e-42        Length:        251
Score:                 583.00          Matches:       110
Percent Similarity:    61.20%          Conservative:  43
Best Local Similarity: 44.00%          Mismatches:    89
Query Match:           32.75%          Indels:        8
DB:                    1               Gaps:          4
```

```
US-09-905-083A-30 (1-969) x KLKE_HUMAN (1-251)

Qy   28  CTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTGGAAACTGCAGGAGAAGAA   87
            |||||  ||||||:::::::  :::::||        |||        :: ||
Db    4  LeuLeuThrAlaLeuGlnValLeuAlaIleAlaMet---------ThrGlnSerGlnGlu   20

Qy   88  GCCCAGGGTGACAAGATTATTGATGGCGCCCCCATGTGCAAGAGAGCTCCACCCATGGCAG  147
            :::     ||||||||||||||||||||||||||||||||||||||||
Db   21  AspGlu---AsnLysIleLeuGlyGlyHisThrCysSerThrArgSerGlnProTrpGln   39

Qy  148  GTGGCCCTGCTCAGTGGC-----AATCAGCTCCACTGCGGAGGCGTCCTGGTCAATGAG  201
            |||||||||||||||:::       ::: |||||||||||||||
Db   40  AlaAlaLeuLeuAlaGlyTyrProArgArgPheLeuCysGlyGlyAlaLeuLeuSerGly   59

Qy  202  CGCTGGGTGCTCACTGCCGCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGT  261
            :::|||||||||||||:::||||||||||||||| ||||| |||||
Db   60  GlnTrpValIleThrAlaAlaHisCysGlyLysMetAsnGluTyrThrValHisLeuGlyLys  79

Qy  262  GATACGGCTG------GGCACAGCAGAGCTCAGAGGATCAAGGACTCGAAGTCATTCCGC  315
            :::              :: ||||| ::|||||||||||| ::  |||
Db   80  HisAsnLeuArgArgTrpGluValThrGlnGlnValLeuArgValValArgGlnValThr   99

Qy  316  CACCCCGGCTACTCCACACAGCCCATGTTAATGACCTCGTCGTGAAGCTCAATAGC  375
            |||||  ||:::::::  ||| ::  |||||||||||||||||:::||
Db  100  HisProAsnTyrAsnSerArgThrHisAspAsnAspLeuMetLeuLeuGlnLeuGln  119

Qy  376  CAGGCCAGGCTGTCATCCATGGTCAAGAAAGTCAGGCTG  435
            |||||:::  ::: |||||||||||||||||||||
Db  120  ProAlaArgIleGlyArgAlaValArgProIleGluValThrGlnAlaCysAlaSerPro  139

Qy  436  GGAACCACCTGACTGTCTCCGGCTGGGCACTACCAGACCGCCAGATGTGACCTTTCC  495
            ||||||:::::::||||||||||||||||||||||||
Db  140  GlyThrSerCysArgValSerGlyTrpGlyTyrThrIleSerSerProIleAlaArgTyrPro  159

Qy  496  TCTGACCTCATGTGCCTGGATGTCAAGCTCATCTCCCCCCAGGACTGCACGGAAGGTTTAC  555
            ::: ||| |||:::::::: :::        ||| |||||:::
Db  160  AlaSerLeuGlnCysValAsnIleAsnIleSerProAspGluValCysGlnLysAlaThrTyr  179

Qy  556  AAGGACTTACTGGAAAATTCCATGCTGTGCGCTGGCATCCCGACTCCCAAGAAAAACGCC  615
            ||:|||||||||||||||||||||
Db  180  ProArgThrIleThrProGlyMetValCysAlaGlyLysValProGlnGlyGlyLysAspSer  199

Qy  616  TGCAATGTGACTCAGGGGGACCGTTGGTGTGCAGAGGTACCCTGCAAGGTTCTGGTGTCC  675
```

```
Db  200  CysGlnGlyAspSerGlyGlyProLeuValCysArgGlyGlnLeuGlnGlyLeuValSer  219
              |||      |||||||||||||||||||||||||||||||||||||||||
Qy  676  TGGGGAACTTCCCTTGCGGCCACCAATGACCCAGAGTCTACACTCAAGTGTCGAAG  735
              ||  || ::|||||||||||||||||||||||||||||||| :: |||||||
Db  220  TrpGlyMetGluArgCysAlaLeuProGlyTyrProGlyValTyrThrAsnLeuCysLys  239

Qy  736  TTCCAAGTGGATAAATGACACCACCAGAA  765
            ::: ||||||         ||:::||
Db  240  TyrArgSerTrpIleGluGluThrMetArg  249
```

```
RESULT 15
Q6B089
ID   Q6B089      PRELIMINARY;      PRT;     251 AA.
AC   Q6B089;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Kallikrein 14, preproprotein.
GN   Name=KLK14;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12277932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RA   Director MGC Project;
RL   Submitted (JUN-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: Belongs to peptidase family S1.
DR   EMBL; BC074905; AAH74905.1; -.
DR   EMBL; BC074904; AAH74904.1; -.
DR   GO; GO:0004263; F:chymotrypsin activity; IEA.
DR   GO; GO:0004295; F:trypsin activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001254; Peptidase_S1.
DR   InterPro; IPR001314; Peptidase_S1A.
DR   InterPro; IPR009003; Pept_Ser_Cys.
DR   Pfam; PF00089; Trypsin; 1.
DR   PRINTS; PR00722; CHYMOTRYPSIN.
DR   SMART; SM00020; Tryp_SPc; 1.
DR   PROSITE; PS50240; TRYPSIN_DOM; 1.
DR   PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
DR   PROSITE; PS00135; TRYPSIN_SER; 1.
KW   Hydrolase; Protease; Serine protease.
SQ   SEQUENCE   251 AA;  27507 MW;  C353A7D8B1960BD2 CRC64;

Alignment Scores:
Pred. No.:     5.97e-42          Length:     251
```

```
Score:                    583.00          Matches:           110
Percent Similarity:       61.20%          Conservative:       43
Best Local Similarity:    44.00%          Mismatches:         89
Query Match:              32.75%          Indels:              8
DB:                       2               Gaps:                4

US-09-905-083A-30 (1-969) x Q6B089 (1-251)

Qy     28 CTTCTCCTGCCCCTGCAGATCTTACTGCTATCCTTAGCCTTGGAAACTGCAGGAGAAGAA  87
          ||||||    |||||||       ||||  :::||::::||
Db      4 LeuLeuThrAlaLeuGlnValLeuAlaIleAlaMet-------ThrArgSerGlnGlu   20

Qy     88 GCCCAGGGTGACAAGATTATTGATGGCGCCCATGTGCAAGAGGCTCCCACCATGGCAG   147
          :::  ::|||||||||    |||  ||| ||||||||
Db     21 AspGlu---AsnLysIleIleGlyGlyTyrThrCysThrArgSerSerGlnProTrpGln   39

Qy    148 GTGGCCCTGCTCAGTGGC------AATCAGTCCACTGCGGAGGGTCCTGGTCAATGAG   201
          |||||||        :::  |||||||   |||:::::
Db     40 AlaAlaLeuLeuAlaGlyProArgArgPheLeuCysGlyGlyAlaLeuLeuSerGly   59

Qy    202 CGCTGGGGTGCTCACTGCGCCCCACTGCAAGATGAATGAGTACACCGTGCACCTGGGCAGT   261
          :::|||||||||:::||||||||||     |||||||
Db     60 GlnTrpValIleThrAlaAlaHisCysGlyArgProIleLeuGlnValAlaLeuGlyLys   79

Qy    262 GATACGGCTG------GGGCACAGGAGAGCTCAGAGGATCAAGGCCTCGAAGTCATTCCGC   315
          |||             |||     :::
Db     80 HisAsnLeuArgArgTrpGluAlaThrGlnGlnValLeuArgValValArgGlnValThr   99

Qy    316 CACCCGGCTACTCCACACAGACCCATGTTAATGACCTCGTGAAGTCAATAGC   375
          |||||     |||||||:::::::||||||||:::||||||
Db    100 HisProAsnTyrAsnSerArgThrHisAspAsnAsnGlnLeuLeuGlnLeuGlnGln   119

Qy    376 CAGGCCAGGCTGTCATCCATGGTGAAGAAAGTCAGGCTGCCCTCCCCGCTGCAACCCCCT   435
          ||||||:::          ::: :::
Db    120 ProAlaArgIleGlyArgAlaValArgProIleGluValThrGlnAlaCysAlaSerPro   139

Qy    436 GGAACCACCTGTACTGTCTCCGGCTGGGGCACTACCACGAGCCCAGATGTGACCTTTCCC   495
          |||:::::||          |||||||||  ::|||||||
Db    140 GlyThrSerCysAlaValSerGlyTrpGlyThrIleSerSerProIleAlaArgTyrPro   159

Qy    496 TCTGACCTCATGTGCGTGGATGTCAAGCTCATCTCCCCCAGGACTGCACGAAGGTTTAC   555
          |||          ::: :::   ||| ||:::|||||||
Db    160 AlaSerLeuGlnCysValAsnIleAsnIleSerProAspGluValCysGlnLysAlaTyr   179

Qy    556 AAGGACTTACTGGAAAAATTCCATGCTGTGCGCTGGCCATCCCCGACTCCAAGAAAAACGCC   615
          :::       ||||||||||||:::||        ||:::
Db    180 ProArgThrIleThrProGlyMetValCysAlaGlyValProGlnGlyGlyLysAspSer   199

Qy    616 TGCAATGGTGACTCAGGGGGACCGTTGGTGTGCAGAGTACCCTGCAAGGTCTGGTGTCC   675
          |||  |||||||||||||||||||||  ||||||||||||||||
Db    200 CysGlnGlyAspSerGlyGlyProLeuValCysArgGlnLeuGlyGlnGlyLeuValSer   219

Qy    676 TGGGGAACTTCCCTTGCCGGCCAACCAATGACCCAGGAGTCTACACTGTCAAGTGTGCAAG   735
          |||||||||||||||  |||||||    |||||||||||||||  :::|||||||
Db    220 TrpGlyMetGluArgGlyCysAlaLeuProGlyTyrProGlyValTyrThrAsnLeuCysLys   239

Qy    736 TTCACCAAGTGGATAAAATGACACCATGAAA   765
          :::       ||||||:::|||||||:::
Db    240 TyrArgSerTrpIleGluGluThrMetArg   249
```

Search completed: July 14, 2005, 17:31:58
Job time : 211.5 secs